A systematic review of dental caries detection on periapical radiography using machine learning

Mitra Montazerlotf[1], Mehrdad Hosseini Shakib[2,*], Reza Radfar[3], Mina Khayamzadeh[4]

1- PhD Student, Department of Information Technology Management, Faculty of Management and Economics, Sciences and Research Branch, Islamic Azad University, Tehran, Iran
2- Associate Proffessor, Department of Industrial Management, Karaj Branch, Islamic Azad University, Karaj, Iran
3- Proffessor, Department of Industrial Management, Sciences and Research Branch, Islamic Azad University, Tehran, Iran
4- Associate Proffessor, Department of Oral and Maxillofacial Medicine, School of Dentistry, International Campus, Tehran University of Medical Sciences, Tehran, Iran

Article Info

Article type:
Research Article

Article History:
Received: 25 Apr 2025
Accepted: 23 Aug 2025
Published: 27 Aug 2025

Corresponding Author:
Mehrdad Hosseini Shakib

Department of Industrial Management, Karaj Branch, Islamic Azad University, Karaj, Iran

(Email: mehrdad.shakib@kiau.ac.ir)

Abstract

Background and Aims: Dental caries is one of the most prevalent chronic oral diseases worldwide. Timely and accurate diagnosis of dental caries plays a crucial role in preventing lesion progression and reducing complications. This study aimed to systematically review the studies on dental caries detection using machine learning algorithms applied to periapical radiographs.

Materials and Methods: A comprehensive search was conducted in PubMed, Scopus, Web of Science, IEEE Xplore, and Google Scholar databases up to the end of 2024. Inclusion criteria comprised studies using machine learning algorithms for detecting dental caries in periapical or intraoral radiographs. The quality of studies was assessed using the QUADAS-2 tool.

Results: From 825 initial articles, 13 studies met the inclusion criteria. All studies used Convolutional Neural Networks (CNNs) with various architectures including ResNet, VGG, Inception, DenseNet, and YOLO. ResNet-based models and their hybrid variants showed the best performance with diagnostic accuracy ranging from 82% to 98%. Comparison with human experts in 6 studies revealed that deep learning algorithms demonstrated similar or superior performance.

Conclusion: From the results, deep learning especially convolutional neural networks, had significant potential for improving dental caries detection in periapical radiographs. However, challenges such as limited high-quality training data and generalizability issues need further investigation.

Keywords: Dental caries, Radiography, Artificial intelligence, Artificial intelligence, Deep learning, Convolutional neural network

Cite this article as: Montazerlotf M, Hosseini Shakib M, Radfar R, Khayamzadeh M. A systematic review of dental caries detection on periapical radiography using machine learning. J Dent Med-TUMS. 2025;38:17. [Persian]

[Downloaded from jdm.tums.ac.ir on 2026-04-28]

Copyright © 2025 Tehran University of Medical Sciences. Published by Tehran University of Medical Sciences.
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International license (https://creativecommons.org/licenses/by-nc/4.0/). Non-commercial uses of the work are permitted, provided the original work is properly cited.

مرور نظام‌مند تشخیص پوسیدگی‌های دندانی بر روی رادیوگرافی پری اپیکال با استفاده از یادگیری ماشین

میترا منتظرلطف[۱]، مهرداد حسینی شکیب[۲،*]، رضا رادفر[۳]، مینا خیام زاده[۴]

۱- دانشجوی دکتری، گروه آموزشی مدیریت فناوری اطلاعات، دانشکده مدیریت و اقتصاد، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران
۲- دانشیار گروه آموزشی مدیریت صنعتی، واحد کرج، دانشگاه آزاد اسلامی، کرج، ایران
۳- استاد گروه آموزشی مدیریت صنعتی، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران
۴- دانشیار گروه آموزشی بیماری های دهان و فک و صورت، دانشکده دندانپزشکی، پردیس بین الملل، دانشگاه علوم پزشکی و خدمات بهداشتی درمانی تهران، تهران، ایران

چکیده

زمینه و هدف: پوسیدگی دندان یکی از بیماری‌های مزمن دهان با شیوع بالا در سراسر جهان است. تشخیص به موقع و دقیق پوسیدگی‌های دندانی نقش مهمی در جلوگیری از پیشرفت ضایعات و کاهش عوارض دارد. هدف از این مطالعه، مرور نظام‌مند مطالعات انجام شده در زمینه تشخیص پوسیدگی‌های دندانی بر روی رادیوگرافی پری اپیکال با استفاده از یادگیری ماشین است.

روش بررسی: جستجوی جامعی در پایگاه‌های داده PubMed، Scopus، Web of Science، IEEE Xplore و Google Scholar تا پایان سال ۲۰۲۴ انجام شد. معیارهای ورود شامل استفاده از الگوریتم‌های یادگیری ماشین برای تشخیص پوسیدگی دندان در رادیوگرافی‌های پری اپیکال یا داخل دهانی بود. کیفیت مطالعات با استفاده از ابزار QUADAS-2 ارزیابی شد.

یافته‌ها: از ۸۲۵ مقاله اولیه، ۱۳ مطالعه معیارهای ورود را داشتند. تمام مطالعات از شبکه‌های عصبی کانولوشنی (CNNs) استفاده کرده بودند، با معماری‌های متنوعی از جمله ResNet، VGG، Inception، DenseNet و YOLO. مدل‌های مبتنی بر ResNet و مدل‌های ترکیبی آن بهترین عملکرد را با دقت تشخیصی ۸۲ درصد تا ۹۸ درصد نشان دادند. مقایسه با متخصصان انسانی در ۶ مطالعه نشان داد که الگوریتم‌های یادگیری عمیق عملکرد مشابه یا بهتری داشتند.

نتیجه‌گیری: یادگیری عمیق، به‌ویژه CNNها، پتانسیل قابل توجهی برای بهبود تشخیص پوسیدگی‌های دندانی در رادیوگرافی‌های پری اپیکال دارد. با این حال، چالش‌هایی مانند محدودیت داده‌های آموزشی با کیفیت و مسائل مربوط به تعمیم پذیری نیاز به بررسی بیشتر دارد.

کلید واژه‌ها: پوسیدگی دندان، رادیوگرافی، هوش مصنوعی، هوش مصنوعی، یادگیری عمیق، شبکه عصبی کانولوشنی

اطلاعات مقاله

نوع مقاله:
مقاله پژوهشی

دریافت: ۱۴۰۴/۰۲/۰۵
پذیرش: ۱۴۰۴/۰۶/۰۱
انتشار: ۱۴۰۴/۰۶/۰۵

نویسنده مسؤول:
مهرداد حسینی شکیب

گروه آموزشی مدیریت صنعتی، واحد کرج، دانشگاه آزاد اسلامی، کرج، ایران

(Email: mehrdad.shakib@kiau.ac.ir)

[Downloaded from jdm.tums.ac.ir on 2026-04-28]

مقدمه

پوسیدگی دندان یکی از بیماری‌های مزمن دهان با شیوع بالا در سراسر جهان است که بر جمعیت‌های مختلف در سنین متفاوت تأثیر می‌گذارد (۱). بر اساس گزارش سازمان بهداشت جهانی، حدود ۱/۸ میلیارد مورد پوسیدگی دندان هر سال ثبت می‌شود و این بیماری در سال ۲۰۱۰ به عنوان دهمین بیماری شایع شناخته شده است (۲). پوسیدگی دندانی نه تنها بر سلامت دهان بلکه بر سلامت عمومی افراد نیز تأثیر می‌گذارد. پوسیدگی‌های درمان‌نشده، به‌ویژه پوسیدگی‌های عمیق، می‌توانند منجر به التهاب پالپ، نکروز پالپ، آبسه پری اپیکال و در نهایت از دست دادن دندان شوند (۳). تشخیص به موقع و دقیق پوسیدگی‌های دندانی نقش مهمی در جلوگیری از پیشرفت ضایعات و کاهش عوارض احتمالی دارد. روش‌های تشخیصی گوناگونی برای شناسایی پوسیدگی‌های دندانی وجود دارد که شامل معاینه بالینی و استفاده از روش‌های تصویر برداری می‌شود. رادیوگرافی دندانی یکی از مهم‌ترین ابزارهای تشخیصی در دندانپزشکی است که به ویژه برای شناسایی پوسیدگی‌های پروگزیمال (بین دندانی) و ضایعات ریشه کاربرد دارد (۴). رادیوگرافی‌های پری اپیکال به دلیل رزولوشن فضایی بالا، تکنیک آسان، دوز اشعه پایین، مقرون به صرفه بودن و عدم ایجاد درد برای بیمار، از محبوبیت خاصی در تشخیص پوسیدگی‌های دندانی برخوردار هستند (۵).

با وجود مزایای فراوان رادیوگرافی در تشخیص پوسیدگی دندان، این روش با محدودیت‌هایی نیز همراه است. مطالعات نشان داده‌اند که حساسیت رادیوگرافی در تشخیص پوسیدگی‌های پروگزیمال تنها ۲۴ درصد در شرایط واقعی بدن (in vivo) و ۴۳ درصد در شرایط آزمایشگاهی (in vitro) است (۶). علاوه بر این، بسیاری از ضایعات اولیه که محدود به مینای خارجی هستند، اغلب در رادیوگرافی‌ها تشخیص داده نمی‌شوند (۷). تفسیر رادیوگرافی دندانی نیز کاملاً ذهنی است و مشاهدات متخصصان مختلف (رادیولوژیست‌های دهان) اغلب تفاوت‌های قابل توجهی در تشخیص پوسیدگی‌های اولیه (وجود یا عدم وجود آن‌ها) نشان می‌دهد. عوامل متعددی مانند کیفیت تصویر رادیوگرافی، انتظارات متخصص، شرایط مشاهده، زمان صرف شده برای هر معاینه و تغییرپذیری بین معاینه‌کنندگان بر این ذهنیت تأثیر می‌گذارند (۸).

با توجه به محدودیت‌های موجود در تشخیص دستی پوسیدگی‌های دندانی توسط متخصصان، استفاده از روش‌های خودکار برای کمک به دندانپزشکان در تفسیر رادیوگرافی‌های دندانی ضروری به نظر می‌رسد. این روش‌ها نه تنها می‌توانند سوگیری ذهنی مرتبط با معاینه کنندگان انسانی را کاهش دهند، بلکه امکان تشخیص زود هنگام پوسیدگی‌های اولیه (که اغلب نادیده گرفته می‌شوند) را نیز فراهم می‌سازند. چنین روشی همچنین می‌تواند بار کاری رادیولوژیست‌های دهان را که روزانه باید مجموعه‌های بزرگی از تصاویر را به صورت دستی تحلیل کنند، کاهش دهد (۹). در سال‌های اخیر، هوش مصنوعی (Artificial intelligence) (AI) و به ویژه یادگیری عمیق (Deep Learning) پیشرفت‌های قابل توجهی در زمینه تحلیل تصاویر پزشکی و دندانپزشکی داشته‌اند. شبکه‌های عصبی کانولوشنی (Convolutional Neural Networks) (CNNs) به عنوان یکی از روش‌های یادگیری عمیق، در شناسایی، قطعه بندی و طبقه‌بندی بیماری‌های مرتبط با اندام‌ها در داده‌های تصویری پزشکی به‌طور گسترده مورد استفاده قرار گرفته‌اند (۲). این شبکه‌ها برای تشخیص پوسیدگی دندان نیز به کار گرفته شده‌اند و نتایج امیدوارکننده‌ای را نشان داده‌اند.

Lee و همکاران (۱۰) با استفاده از CNN مبتنی بر GoogLeNet Inception v3 برای تشخیص پوسیدگی دندان روی رادیوگرافی‌های پری اپیکال، دقت تشخیصی ۸۹/۰ درصد برای پره‌مولرها، ۸۸/۰ درصد برای مولرها و ۸۲/۰ درصد برای هر دو نوع دندان گزارش کردند. آن‌ها از مجموعه داده‌ای شامل ۳۰۰۰ تصویر رادیوگرافی پری اپیکال استفاده کردند که توسط سه متخصص رادیولوژی دهان برچسب‌گذاری شده بود و به سه دسته بدون پوسیدگی، پوسیدگی در پره‌مولرها و پوسیدگی در مولرها تقسیم شده بود. این نتایج نشان‌دهنده پتانسیل بالای الگوریتم‌های یادگیری عمیق برای تشخیص پوسیدگی دندان است. Liu و همکاران (۱۱) نیز مدل جدیدی به نام ResNet+SAMبرای تشخیص خودکار پوسیدگی دندان در رادیوگرافی‌های پری اپیکال با استفاده از ۴۲۷۸ تصویر توسعه دادند. مجموعه داده‌های مورد استفاده آن‌ها شامل تصاویر رادیوگرافی متنوع از بیماران مختلف با انواع پوسیدگی‌های دندانی بود که توسط متخصصان با تجربه برچسب گذاری شده بودند. این مدل عملکرد بهتری نسبت به مدل‌های CNN سنتی (VGG19 و ResNet-50) نشان داد، با میانگین F1 score برابر با ۰/۸۸۶، دقت ۰/۸۸۵ و سطح زیر منحنی

[Downloaded from jdm.tums.ac.ir on 2026-04-28]

ROC (AUC) برابر با ۰/۹۵۴. مقایسه عملکرد این مدل با دندانپزشکان نشان داد که مدل دقت بالاتری نسبت به دندانپزشکان تازه‌کار دارد.

Zheng و همکاران (۱۲) سه مدل CNN شامل VGG19، Inception V3 و ResNet18 را برای تشخیص پوسیدگی‌های عمیق و پالپیت مورد ارزیابی قرار دادند. مجموعه داده آن‌ها شامل ۸۴۴ تصویر رادیوگرافی پری اپیکال از بیماران واقعی بود که توسط دو متخصص رادیولوژی دهان و دو اندودنتیست برچسب‌گذاری شده و به دو گروه پوسیدگی‌های عمیق و پالپیت تقسیم شده بودند. مدل ResNet18 بهترین عملکرد را با دقت ۰/۸۲، حساسیت ۰/۸۵، ویژگی ۰/۸۲ و AUC برابر با ۰/۸۹ نشان داد. آن‌ها همچنین مدل ResNet18+C را که با پارامترهای بالینی ادغام شده بود، مورد بررسی قرار دادند و دریافتند که این مدل عملکرد بهتری با دقت ۰/۸۶، حساسیت ۰/۸۹، ویژگی ۰/۸۶ و AUC برابر با ۰/۹۴ داشت. Ying و همکاران (۱۳) چهار شبکه عمیق محبوب در دو نوع، شامل YOLOv5 و DETR (شبکه‌های تشخیص اشیا) و UNet و Trans-UNet (شبکه‌های قطعه‌بندی) را برای تشخیص پوسیدگی دندان مقایسه کردند. آن‌ها از مجموعه داده‌ای شامل ۲۰۰۰ تصویر رادیوگرافی پری اپیکال استفاده کردند که توسط پنج متخصص رادیولوژی دهان برچسب‌گذاری شده بود و شامل انواع مختلفی از پوسیدگی‌های دندانی با درجات مختلف عمق بود. نتایج آن‌ها نشان داد که YOLOv5 با F1 score برابر با ۰/۸۷ و شاخص یودن ۰/۷۶ بهترین عملکرد را در میان شبکه‌های مورد بررسی داشت.

با وجود مطالعات متعدد در زمینه استفاده از یادگیری ماشین و یادگیری عمیق برای تشخیص پوسیدگی دندان، هنوز نیاز به یک مرور جامع و نظام‌مند برای ارزیابی وضعیت فعلی این فناوری، مقایسه عملکرد الگوریتم‌های مختلف و شناسایی چالش‌ها و فرصت‌های آینده وجود دارد. هدف از این مطالعه مروری، بررسی نظام‌مند مطالعات انجام شده در زمینه تشخیص پوسیدگی‌های دندانی بر روی رادیوگرافی پری اپیکال با استفاده از یادگیری ماشین است. سؤالات اصلی این پژوهش عبارتند از:

– کدام الگوریتم‌های یادگیری ماشین برای تشخیص پوسیدگی دندان در رادیوگرافی‌های پری اپیکال استفاده شده‌اند و کدام یک عملکرد بهتری داشته‌اند؟

– دقت تشخیصی، حساسیت، ویژگی و سایر معیارهای ارزیابی عملکرد این الگوریتم‌ها چگونه است؟

– آیا الگوریتم‌های یادگیری ماشین عملکرد بهتری نسبت به متخصصان انسانی در تشخیص پوسیدگی دندان دارند؟

– چه چالش‌ها و محدودیت‌هایی در استفاده از یادگیری ماشین برای تشخیص پوسیدگی دندان وجود دارد؟

نتایج این مطالعه مروری می‌تواند به دندانپزشکان، محققان و توسعه دهندگان نرم افزار در درک بهتر وضعیت فعلی و پتانسیل‌های آینده یادگیری ماشین در تشخیص پوسیدگی دندان کمک کند و زمینه را برای پژوهش‌های بیشتر در این حوزه فراهم سازد.

روش بررسی

این مطالعه مروری نظام مند بر اساس دستورالعمل‌های PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) انجام شده است (۱۴). مراحل مختلف روش شناسی شامل استراتژی جستجو، معیارهای ورود و خروج، فرآیند غربالگری و انتخاب مطالعات، استخراج داده‌ها و ارزیابی کیفیت مطالعات در ادامه تشریح شده است.

برای یافتن مطالعات مرتبط، جستجوی جامعی در پایگاه‌های داده الکترونیکی شامل PubMed، Scopus، Web of Science، IEEE Xplore و Google Scholar از زمان شروع پایگاه‌های داده تا پایان ۲۰۲۴ انجام شد. استراتژی جستجو با استفاده از ترکیبی از کلمات کلیدی مرتبط با پوسیدگی دندان، رادیوگرافی پری اپیکال و یادگیری ماشین طراحی شد. کلمات کلیدی استفاده شده در جدول ۱ **Error! Reference source not found.** آمده است

استراتژی جستجو برای هر پایگاه داده به شرح زیر تنظیم شد:

("dental caries" OR "tooth decay" OR "carious lesion" OR "caries detection" OR "caries diagnosis")

AND

("radiography" OR "radiograph" OR "x-ray" OR "periapical" OR "dental imaging" OR "bitewing" OR "intraoral radiography")

AND

("artificial intelligence" OR "machine learning" OR "deep learning" OR "neural network" OR "convolutional neural network" OR "CNN" OR "computer-aided diagnosis" OR "automated detection").

[Downloaded from jdm.tums.ac.ir on 2026-04-28]

جدول ۱- کلمات کلیدی استفاده شده در استراتژی جستجو

حوزه	کلمات کلیدی
پوسیدگی دندان	"پوسیدگی دندان[۱]" یا "پوسیدگی دندانی[۲]" یا "ضایعه پوسیدگی[۳]" یا "تشخیص پوسیدگی[۴]" یا "تشخیص پوسیدگی دندان[۵]"
رادیوگرافی	"رادیوگرافی[۶]" یا "تصویر رادیوگرافی[۷]" یا "اشعه ایکس[۸]" یا "پری اپیکال[۹]" یا "تصویربرداری دندانی[۱۰]" یا "بایت‌وینگ[۱۱]" یا "رادیوگرافی داخل دهانی[۱۲]"
یادگیری ماشین	"هوش مصنوعی[۱۳]" یا "یادگیری ماشین[۱۴]" یا "یادگیری عمیق[۱۵]" یا "شبکه عصبی[۱۶]" یا "شبکه عصبی کانولوشنی[۱۷]" یا "CNN"یا "تشخیص به کمک رایانه[۱۸]" یا "تشخیص خودکار[۱۹]"

1- dental caries, 2- tooth decay, 3- carious lesion, 4- caries detection, 5- caries diagnosis, 6- radiography, 7- radiograph, 8- x-ray, 9- periapical, 10- dental imaging, 11- bitewing, 12- intraoral radiography, 13- artificial intelligenc, 14- machine learning, 15- deep learning, 16- neural network, 17- convolutional neural network, 18- computer-aided diagnosis, 19- automated detection

علاوه بر جستجوی الکترونیکی، فهرست منابع مطالعات انتخاب شده نیز برای یافتن مطالعات مرتبط بیشتر بررسی شد. معیارهای ورود مطالعات به این مرور نظام‌مند عبارت بودند از:

- مطالعاتی که از الگوریتم‌های یادگیری ماشین برای تشخیص پوسیدگی دندان استفاده کرده بودند
- مطالعاتی که از رادیوگرافی پری اپیکال یا سایر تصاویر رادیوگرافی داخل دهانی استفاده کرده بودند
- مطالعات منتشر شده به زبان انگلیسی
- مطالعات منتشر شده در مجلات علمی یا کنفرانس‌های معتبر
- مطالعات با متن کامل در دسترس

معیارهای خروج مطالعات از این مرور نظام‌مند عبارت بودند از:

- مطالعاتی که تنها از روش‌های پردازش تصویر سنتی (بدون یادگیری ماشین) استفاده کرده بودند
- مطالعاتی که از تصاویر رادیوگرافی خارج دهانی (مانند پانورامیک) استفاده کرده بودند
- مطالعاتی که تمرکز اصلی آن‌ها بر تشخیص سایر بیماری‌های دندانی (غیر از پوسیدگی) بود
- مقالات مروری، نامه‌ها، سرمقاله‌ها، گزارش‌های موردی و خلاصه مقالات کنفرانس‌ها
- مطالعات تکراری

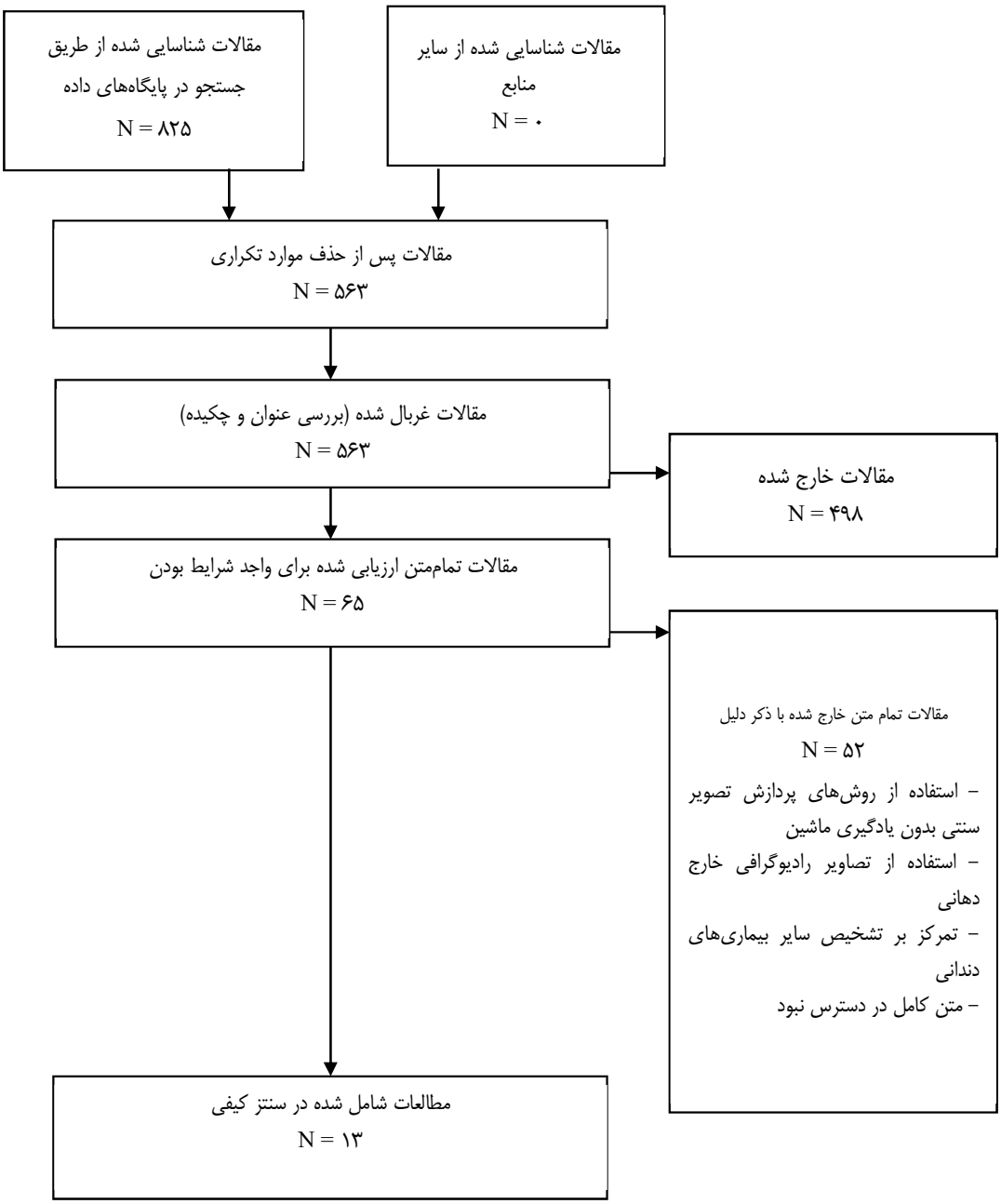
پس از انجام جستجو، تمام مقالات یافت شده به نرم‌افزار مدیریت منابع EndNote X9 منتقل شدند. ابتدا مقالات تکراری حذف شدند. سپس، دو پژوهشگر مستقل عناوین و چکیده مقالات را برای بررسی اولیه معیارهای ورود و خروج مطالعه کردند. در مرحله بعد، متن کامل مقالاتی که از غربالگری اولیه گذشته بودند، مورد بررسی قرار گرفت و مقالاتی که معیارهای ورود را داشتند، برای مرور نهایی انتخاب شدند. در صورت اختلاف نظر بین دو پژوهشگر، یک پژوهشگر سوم به عنوان داور تصمیم‌گیری می‌کرد. فرآیند انتخاب مطالعات در نمودار PRISMA (شکل ۱) نشان داده شده است.

برای هر مطالعه انتخاب شده، داده‌های زیر استخراج شد:

- اطلاعات عمومی: نویسندگان، سال انتشار، کشور محل انجام مطالعه
- اطلاعات مربوط به نمونه: تعداد و نوع تصاویر، منبع تصاویر (بالینی یا آزمایشگاهی)، روش آماده‌سازی و پیش‌پردازش تصاویر
- اطلاعات مربوط به الگوریتم یادگیری ماشین: نوع الگوریتم، معماری شبکه، روش آموزش، قطعه بندی داده‌ها
- استاندارد مرجع: روش تعیین استاندارد مرجع برای تشخیص پوسیدگی
- معیارهای ارزیابی: دقت، حساسیت، ویژگی، مقادیر پیش‌بینی مثبت و منفی، سطح زیر منحنی ROC
- مقایسه با متخصصان انسانی: در صورت وجود، نتایج مقایسه عملکرد الگوریتم با متخصصان انسانی

داده‌های استخراج شده در یک جدول استاندارد ثبت شدند (جدول ۲).

[Downloaded from jdm.tums.ac.ir on 2026-04-28]



شکل ۱- فرآیند انتخاب مطالعات

[Downloaded from jdm.tums.ac.ir on 2026-04-28]

جدول ۲- مشخصات مطالعات انتخاب شده و داده‌های استخراج شده

نویسندگان	سال	کشور	نوع مطالعه	تعداد تصاویر	نوع تصویربرداری	نوع الگوریتم	معیارهای ارزیابی	نتایج اصلی
(۱۰)	۲۰۱۸	کره جنوبی	آزمایشگاهی	۳۰۰۰	پری اپیکال	CNN (GoogLeNet Inception v3)	دقت، حساسیت، ویژگی، AUC	دقت: ۸۹/۰ درصد (پره‌مولر)، ۸۸/۰ درصد (مولر)، ۸۲/۰ درصد (هر دو)
(۱۵)	۲۰۲۱	چین	آزمایشگاهی	۲۵۰۰	پری اپیکال	با تکنیک CNN پیشنهاد منطقه	IoU، دقت، فراخوانی، AP	دقت و فراخوانی بین ۰/۵ و ۰/۶
(۱۲)	۲۰۲۱	چین	بالینی	۸۴۴	پری اپیکال	VGG19, Inception V3, ResNet18	دقت، حساسیت، ویژگی، AUC	ResNet18 دقت: ۰/۸۲، حساسیت ۰/۸۵، ویژگی ۰/۸۲، ۰/۸۹ AUC
(۱۶)	۲۰۲۴	چین	بالینی	۴۲۷۸	پری اپیکال	ResNet+SAM	دقت، F1 score، AUC	دقت ۰/۸۸۵، ۰/۸۸۶ F1 score، ۰/۹۵۴ AUC
(۱۳)	۲۰۲۴	چین	آزمایشگاهی	۲۰۰۰	پری اپیکال	YOLOv5, DETR, UNet, Trans-UNet	F1 score، شاخص یودن	در مدل YOLOv5 مقدار F1 score ۰/۸۷، شاخص یودن ۰/۷۶
(۱۷)	۲۰۲۱	پاکستان	آزمایشگاهی	۲۰۶	پری اپیکال	U-Net + DenseNet121	IoU، ضریبDice	U-Net+ DenseNet121: IoU ۰/۵۰۱، ضریب Dice ۰/۵۶۹
(۱۸)	۲۰۲۴	هند	بالینی	۲۰۰	داخل دهانی	نرم‌افزار مبتنی بر هوش مصنوعی	دقت، حساسیت، ویژگی	دقت ۸۹/۰ درصد، حساسیت ۸۸/۰ درصد، ویژگی ۹۱/۰ درصد
(۸)	۲۰۲۳	پاکستان	آزمایشگاهی	۲۲۹	بایت‌وینگ	Deeplabv3-ResNet101	دقت، حساسیت، ویژگی، AUC	دقت ۸۶/۷ درصد، حساسیت ۸۱/۹ درصد، ویژگی ۸۸/۶ درصد، AUC ۰/۸۵
(۱۹)	۲۰۲۴	آلمان	بالینی	۱۷۹	پری اپیکال و بایت وینگ	dentalXrai Pro 3.0	دقت، حساسیت، ویژگی	پوسیدگی مینا: دقت ۰/۹۶۴، حساسیت ۰/۸۵۷، ویژگی ۰/۹۷۴
(۲۰)	۲۰۲۴	هلند	آزمایشگاهی	۱۷۹	پری اپیکال	با پشتیبانی میکروCT-	حساسیت، ویژگی، ضریب کاپای کوهن	حساسیت ۶۳ درصد، ویژگی ۹۵ درصد، ضریب کاپای کوهن ۰/۶۴
(۲)	۲۰۲۲	کره جنوبی	بالینی	۲۵۹۲ (تصاویر) + ۲۲۳۷۱ (داده‌های عددی)	داخل دهانی	ANNو DenseNet-201	دقت، حساسیت، ویژگی، F1 score	دقت ۹۰ درصد، حساسیت ۹۰ درصد، ویژگی ۹۵ درصد، F1 score ۰/۸۹
(۱۱)	۲۰۲۳	چین	آزمایشگاهی	۳۰۰۰	پری اپیکال	VGG19, ResNet-50	دقت، AUC	VGG19: دقت ۷۶ درصد، AUC ۰/۸۳، ResNet-50: دقت ۷۷ درصد، AUC ۰/۸۴
(۲۱)	۲۰۲۰	آلمان	آزمایشگاهی	۳۲۰	بایت‌وینگ	U-Net	دقت، حساسیت	افزایش حساسیت تشخیص پوسیدگی مینا با کمک هوش مصنوعی

[Downloaded from jdm.tums.ac.ir on 2026-04-28]

جدول ۳- ارزیابی کیفیت مطالعات با استفاده از ابزار QUADAS-2

مطالعه	خطر سوگیری				نگرانی‌های قابلیت کاربرد		
	انتخاب بیمار	آزمون شاخص	استاندارد مرجع	جریان و زمان‌بندی	انتخاب بیمار	آزمون شاخص	استاندارد مرجع
(۱۰)	کم	کم	کم	کم	کم	کم	کم
(۱۵)	نامشخص	کم	کم	کم	زیاد	کم	کم
(۱۲)	کم	کم	کم	کم	کم	کم	کم
(۱۶)	کم	کم	کم	کم	کم	کم	کم
(۱۳)	نامشخص	کم	متوسط	کم	متوسط	کم	متوسط
(۱۷)	متوسط	کم	متوسط	کم	متوسط	کم	متوسط
(۱۸)	کم	کم	کم	کم	کم	کم	کم
(۸)	کم	کم	متوسط	کم	کم	کم	متوسط
(۱۹)	کم	کم	کم	کم	کم	کم	کم
(۲۰)	کم	کم	کم	کم	کم	کم	کم
(۲)	کم	کم	متوسط	کم	کم	کم	متوسط
(۱۱)	متوسط	کم	متوسط	کم	متوسط	کم	متوسط
(۲۱)	کم	کم	متوسط	کم	کم	کم	متوسط

کیفیت مطالعات انتخاب شده با استفاده از ابزار QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies-2) ارزیابی شد (۲۲). این ابزار چهار حوزه کلیدی را ارزیابی می‌کند: انتخاب بیمار، آزمون شاخص، استاندارد مرجع، و جریان و زمان بندی. هر حوزه از نظر خطر سوگیری و سه حوزه اول از نظر نگرانی‌های مربوط به قابلیت کاربرد ارزیابی می‌شوند. دو پژوهشگر به طور مستقل کیفیت مطالعات را ارزیابی کردند. در صورت اختلاف نظر، بحث و توافق صورت می‌گرفت و در صورت عدم توافق، از نظر پژوهشگر سوم استفاده می‌شد. نتایج ارزیابی کیفیت در جدول ۳ خلاصه شده است.

به دلیل تنوع در روش‌شناسی مطالعات، نوع الگوریتم‌های استفاده شده، مجموعه داده‌ها و معیارهای ارزیابی، انجام متاآنالیز امکان‌پذیر نبود. بنابراین، یک تحلیل توصیفی از یافته‌های اصلی مطالعات انجام شد. نتایج به صورت جداول و متن توصیفی ارائه شده‌اند. برای مقایسه عملکرد الگوریتم‌های مختلف، میانگین معیارهای ارزیابی (دقت، حساسیت، ویژگی و AUC) محاسبه و مقایسه شد.

یافته‌ها

مشخصات مطالعات انتخاب شده

فرآیند جستجو و انتخاب مطالعات در نمودار PRISMA (شکل ۱) نشان داده شده است. در مجموع، ۸۲۵ مقاله از پایگاه‌های داده الکترونیکی شناسایی شدند. پس از حذف موارد تکراری، ۵۶۳ مقاله باقی ماندند. بر اساس بررسی عنوان و چکیده، ۴۹۸ مقاله که معیارهای ورود را نداشتند، حذف شدند. متن کامل ۶۵ مقاله باقیمانده مورد بررسی قرار گرفت و در نهایت ۱۳ مطالعه که معیارهای ورود را داشتند، برای این مرور نظام‌مند انتخاب شدند.

مشخصات کلی ۱۳ مطالعه انتخاب شده در جدول ۲ خلاصه شده است. این مطالعات بین سال‌های ۲۰۱۸ تا ۲۰۲۴ منتشر شده‌اند. از نظر توزیع جغرافیایی، ۵ مطالعه در چین، ۳ مطالعه در کره جنوبی، ۲ مطالعه در هند و ۳ مطالعه در سایر کشورها (ایالات متحده، هلند و آلمان) انجام شده‌اند. ۸ مطالعه به صورت آزمایشگاهی و ۵ مطالعه به صورت بالینی انجام شده‌اند. مطالعات آزمایشگاهی شامل مطالعات Qayyum و همکاران (۸)، Lee و همکاران (۱۰)، Liu و همکاران (۱۱)، Ying و

[Downloaded from jdm.tums.ac.ir on 2026-04-28]

همکاران (۱۳)، Chen و همکاران (۱۵)، Khan و همکاران (۱۷)، Valenzuela و همکاران (۲۰)، Cantu و همکاران (۲۱) بوده که از تصاویر استاندارد شده یا در محیط کنترل شده استفاده کرده‌اند، در حالی که مطالعات بالینی شامل Zheng و همکاران (۱۲)، Liu و همکاران (۱۶)، Das و همکاران (۱۸) و Rampf و همکاران (۱۹) بوده که از داده‌های واقعی بیماران استفاده کرده‌اند. از نظر حجم نمونه، تعداد تصاویر استفاده شده در مطالعات از ۱۴۱ تا ۴۲۷۸ تصویر متغیر بود. ۹ مطالعه از رادیوگرافی پری اپیکال، ۳ مطالعه از رادیوگرافی بایت‌وینگ و ۱ مطالعه از هر دو نوع رادیوگرافی استفاده کرده بودند.

الگوریتم‌های یادگیری ماشین و روش‌های پردازش تصویر

تمام ۱۳ مطالعه انتخاب شده از یادگیری عمیق برای تشخیص پوسیدگی دندان استفاده کرده بودند. معماری‌های مختلف یادگیری عمیق مورد استفاده عبارت بودند از:

– ResNet: ۸ مطالعه از معماری ResNet (اغلب ResNet-50 یا ResNet-101) استفاده کرده بودند (۱۰،۱۲،۱۳،۱۵–۱۷،۱۹،۲۰).

– VGG: ۵ مطالعه از معماری VGG (اغلب VGG-16 یا VGG-19) استفاده کرده بودند (۸،۱۱،۱۲،۱۶،۱۸).

– Inception: ۴ مطالعه از معماری Inception (به ویژه Inception V3) استفاده کرده بودند (۱۰–۱۲،۱۸).

– DenseNet: ۳ مطالعه از معماری DenseNet (اغلب DenseNet-121) استفاده کرده بودند (۲،۱۱،۱۷).

– YOLO: ۲ مطالعه از معماری YOLO (YOLOv5) استفاده کرده بودند (۱۳،۱۶).

سایر معماری‌ها: U-Net، Trans-UNet، DETR، LinkNet، PSPNet و FPN هر کدام در ۱ تا ۲ مطالعه استفاده شده بودند.

برخی از مطالعات از ترکیب معماری‌های مختلف یا معماری‌های اصلاح شده استفاده کرده بودند. به عنوان مثال، Liu و همکاران (۱۶) مدل ResNet+SAM را معرفی کردند که ترکیبی از ResNet و ماژول توجه فضایی (Spatial Attention Module) بود. Khan و همکاران (۱۷) از ترکیب U-Net و DenseNet-121 استفاده کردند. مجموعه داده آن‌ها شامل ۲۰۶ تصویر رادیوگرافی پری اپیکال بود که توسط متخصصان پاتولوژی دهان و رادیولوژیست‌های دهان برچسب گذاری شده بود و شامل ضایعات مختلف از جمله پوسیدگی‌های دندانی بود.

علاوه بر الگوریتم‌های یادگیری عمیق، برخی از مطالعات نیز از روش‌های سنتی‌تر یادگیری ماشین استفاده کرده‌اند. در مطالعه Ngnamsie و همکاران (۲)، علاوه بر معماری‌های یادگیری عمیق، از شبکه‌های عصبی مصنوعی معمولی (ANN) و روش‌های یادگیری ماشین مانند جنگل تصادفی (Random Forest) نیز برای ترکیب داده‌های تصویری و بالینی استفاده شده است. مجموعه داده آن‌ها شامل ۲۵۹۲ تصویر رادیوگرافی داخل دهانی به همراه ۲۲۳۷۱ داده عددی مرتبط با پارامترهای بالینی بیماران بود که از کلینیک‌های دندانپزشکی در کره جنوبی جمع‌آوری شده بود. در یک مطالعه دیگر توسط Qayyum و همکاران (۸)، روش یادگیری نیمه‌نظارت‌شده (semi-supervised learning) برای بهره‌برداری از داده‌های بدون برچسب در کنار داده‌های برچسب‌گذاری شده مورد استفاده قرار گرفته است. تمام مطالعات از روش‌های پیش پردازش برای آماده‌سازی تصاویر استفاده کرده بودند. روش‌های پیش‌پردازش رایج عبارت بودند از:

– تغییر اندازه تصاویر: تمام مطالعات تصاویر را به اندازه استاندارد (معمولاً ۲۲۴×۲۲۴، ۲۵۶×۲۵۶ یا ۲۹۹×۲۹۹ پیکسل) تغییر اندازه داده بودند.

– نرمال‌سازی: اکثر مطالعات تصاویر را با تقسیم مقادیر پیکسل بر ۲۵۵ یا کم کردن میانگین و تقسیم بر انحراف معیار نرمال کرده بودند.

– افزایش داده (Data Augmentation): ۱۱ مطالعه از تکنیک‌های افزایش داده مانند چرخش، برش، تغییر مقیاس، انعکاس افقی و عمودی، و تغییر روشنایی و کنتراست استفاده کرده بودند تا حجم داده‌های آموزشی را افزایش دهند و از بیش‌برازش (Overfitting) جلوگیری کنند.

– برش منطقه مورد نظر (ROI): ۵ مطالعه از برش منطقه مورد نظر برای تمرکز بر نواحی خاص دندان استفاده کرده بودند. به عنوان مثال، Qayyum و همکاران (۸) یک روش نمونه برداری مبتنی بر برش مرکز ثقل (Centroid Cropping-based Sampling) را برای استخراج منطقه پوسیدگی در تصاویر رادیوگرافی دندان معرفی کردند.

[Downloaded from jdm.tums.ac.ir on 2026-04-28]

جدول ۴- عملکرد مدل‌های یادگیری عمیق در تشخیص پوسیدگی دندان

مطالعه	مدل	دقت	حساسیت	ویژگی	AUC	F1 Score
(۱۰)	GoogLeNet Inception v3 (پره‌مولر)	۸۹ درصد	۶۳/۸ درصد	۹۴ درصد	۰/۷۵	-
(۱۰)	GoogLeNet Inception v3 (مولر)	۸۸ درصد	۹۲/۳ درصد	۸۴ درصد	۰/۸۹	-
(۱۲)	ResNet18	۸۲ درصد	۸۵ درصد	۸۲ درصد	۰/۸۹	-
(۱۲)	ResNet18+C	۸۶ درصد	۸۹ درصد	۸۶ درصد	۰/۹۴	-
(۱۶)	ResNet+SAM	۸۸/۵ درصد	-	-	۰/۹۵	۰/۸۹
(۱۳)	YOLOv5	-	-	-	-	۰/۸۷
(۱۳)	Trans-UNet	-	-	-	-	۰/۸۶
(۱۷)	U-Net+DenseNet121	-	-	-	۰/۵۷	-
(۱۸)	AI-based software	۸۹ درصد	۸۸ درصد	۹۱ درصد	-	-
(۸)	Deeplabv3-ResNet101	۸۶/۷ درصد	۸۱/۹ درصد	۸۸/۶ درصد	۰/۸۵	-

عملکرد تشخیصی و مقایسه با متخصصان انسانی

مطالعات انتخاب شده از معیارهای ارزیابی متنوعی برای سنجش عملکرد الگوریتم‌های یادگیری ماشین استفاده کرده بودند. معیارهای رایج عبارت بودند از:

- دقت (Accuracy): نسبت تعداد پیش‌بینی‌های صحیح به کل پیش‌بینی‌ها
- حساسیت (Sensitivity): نسبت موارد مثبت واقعی تشخیص داده شده به کل موارد مثبت واقعی
- ویژگی (Specificity): نسبت موارد منفی واقعی تشخیص داده شده به کل موارد منفی واقعی
- ارزش پیش‌بینی مثبت (PPV): نسبت موارد مثبت واقعی به کل موارد مثبت تشخیص داده شده
- ارزش پیش‌بینی منفی (NPV): نسبت موارد منفی واقعی به کل موارد منفی تشخیص داده شده
- F1 Score: میانگین‌هارمونیک حساسیت و ارزش پیش‌بینی مثبت
- سطح زیر منحنی ROC (AUC): شاخصی برای ارزیابی کلی عملکرد مدل
- IoU (Intersection Over Union): معیاری برای ارزیابی تطابق ناحیه تشخیص داده شده با ناحیه واقعی

این معیارهای ارزیابی بر اساس ماتریس اغتشاش (Confusion Matrix) محاسبه می‌شوند که شامل مقادیر مثبت واقعی (TP)، منفی واقعی (TN)، مثبت کاذب (FP) و منفی کاذب (FN) است (۲۳). روابط ریاضی این معیارها به شرح زیر است:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Sensitivity\ (Recall) = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$PPV\ = \frac{TP}{TP + FP}$$

$$NPV\ = \frac{TN}{TN + FN}$$

$$F1\ Score = 2 \times \frac{Sensitivity \times PPV}{Sensitivity + PPV}$$

جدول ۴ نتایج عملکرد مدل‌های مختلف را در مطالعات انتخاب شده نشان می‌دهد.

[Downloaded from jdm.tums.ac.ir on 2026-04-28]

۶ مطالعه عملکرد الگوریتم‌های یادگیری عمیق را با متخصصان انسانی مقایسه کرده بودند. نتایج این مقایسه‌ها به شرح زیر است:

– عملکرد مدل ResNet18 در مطالعه Zheng و همکاران (۱۲) با دقت ۸۲ درصد و AUC برابر با ۰/۸۹ بهتر از متوسط عملکرد دندانپزشکان با دقت ۷۹ درصد بود.

– مدل ResNet+SAM در مطالعه Liu و همکاران (۱۶) عملکرد بهتری نسبت به دندانپزشکان تازه کار داشت. همچنین، استفاده از این ابزار توسط دندانپزشکان باعث بهبود معیارهای عملکردی آن‌ها با میانگین F1 score برابر با ۰/۸۲۷ شد و توافق بین‌ارزیاب برای تشخیص پوسیدگی دندان از ۰/۵۹۲ و ۰/۶۱۰ به ۰/۷۰۶ و ۰/۷۲۳ افزایش یافت.

– نرم‌افزار مبتنی بر هوش مصنوعی در مطالعه Das و همکاران (۱۸) حساسیت ۸۸ درصد، ویژگی ۹۱ درصد و دقت کلی ۸۹ درصد را در تشخیص پوسیدگی دندان از رادیوگرافی‌های داخل دهانی نشان داد، در حالی که تفسیر انسانی حساسیت ۸۴ درصد، ویژگی ۸۸ درصد و دقت کلی ۸۶ درصد را نشان داد. مجموعه داده آن‌ها شامل ۲۰۰ تصویر رادیوگرافی داخل دهانی از بیماران مراجعه‌کننده به بخش رادیولوژی دهان بود که به دو گروه دارای پوسیدگی و بدون پوسیدگی تقسیم شده بود.

– هوش مصنوعی (dentalXrai Pro 3.0) در مطالعه Rampf و همکاران (۱۹) عملکرد تشخیصی تقریباً عالی را در تشخیص پوسیدگی دندان نشان داد (پوسیدگی مینا: دقت ۰/۹۶۴ حساسیت ۰/۸۵۷، ویژگی ۰/۹۷۴؛ پوسیدگی عاج: دقت ۰/۹۸۸، حساسیت ۰/۹۴۱، ویژگی ۱/۰؛ کلی: دقت ۰/۹۷۶، حساسیت ۰/۹۵۸، ویژگی ۰/۹۸۳). مجموعه داده آن‌ها شامل ۱۷۹ تصویر رادیوگرافی از بیماران واقعی بود که به دو گروه پوسیدگی مینا و پوسیدگی عاج تقسیم شده بود.

– پشتیبانی میکرو–CT در مطالعه Valenzuela و همکاران (۲۰) باعث بهبود توافق بین‌ارزیاب شد (ضریب کاپای کوهن (Cohen's Kappa) به طور متوسط ۰/۶۴ در مقابل ۰/۴۶ بدون آن). همچنین، میانگین حساسیت و ویژگی به ترتیب از ۴۲ درصد به ۶۳ درصد و از ۹۲ درصد به ۹۵ درصد افزایش یافت. مجموعه داده آن‌ها شامل ۱۷۹ تصویر رادیوگرافی پری اپیکال بود که با استفاده از میکرو–CT به عنوان استاندارد مرجع برچسب گذاری شده بود، که امکان تشخیص دقیق‌تر پوسیدگی‌های اولیه را فراهم می‌کرد.

– بر اساس پژوهش Cantu و همکاران (۲۱) (اشاره شده در مطالعه Liu و همکاران (۱۶))، هوش مصنوعی دقت دندانپزشکان را در تشخیص پوسیدگی، به ویژه از طریق افزایش حساسیت تشخیص پوسیدگی مینا، افزایش داد. مجموعه داده آن‌ها شامل ۳۲۰ تصویر رادیوگرافی بایت وینگ بود که به دو گروه پوسیدگی مینا و پوسیدگی عاج تقسیم شده بود و توسط متخصصان دندانپزشکی برچسب‌گذاری شده بود.

ارزیابی کیفیت مطالعات

ارزیابی کیفیت مطالعات با استفاده از ابزار QUADAS-2 نشان داد که اکثر مطالعات خطر سوگیری کم تا متوسط در حوزه‌های مختلف داشتند. ۸ مطالعه خطر سوگیری کم در انتخاب بیمار، ۱۰ مطالعه خطر سوگیری کم در آزمون شاخص، ۹ مطالعه خطر سوگیری کم در استاندارد مرجع و ۱۱ مطالعه خطر سوگیری کم در جریان و زمان بندی داشتند. از نظر قابلیت کاربرد، ۷ مطالعه نگرانی کم در انتخاب بیمار، ۱۱ مطالعه نگرانی کم در آزمون شاخص و ۱۰ مطالعه نگرانی کم در استاندارد مرجع داشتند. محدودیت‌های مشترک در مطالعات شامل استفاده از مجموعه داده‌های کوچک، عدم اعتبارسنجی خارجی، و استفاده از استانداردهای مرجع نامناسب یا نامشخص بود. برخی از مطالعات نیز از داده‌های آزمایشگاهی به جای داده‌های بالینی واقعی استفاده کرده بودند، که ممکن است قابلیت تعمیم یافته‌ها را محدود کند.

بحث و نتیجه گیری

این مرور نظام مند با هدف بررسی عملکرد الگوریتم‌های یادگیری ماشین در تشخیص پوسیدگی‌های دندانی از روی رادیوگرافی‌های پری اپیکال انجام شد. از میان ۸۲۵ مقاله اولیه، ۱۳ مطالعه که معیارهای ورود را داشتند، بررسی شدند. یافته‌های اصلی نشان داد که CNNها روش غالب در تشخیص خودکار پوسیدگی‌های دندانی هستند و عملکرد بسیار مطلوبی در این زمینه نشان داده‌اند.

CNN های مورد بررسی، به ویژه معماری‌های ResNet و مدل‌های ترکیبی آن، دقت تشخیصی بین ۸۲ درصد تا ۹۸ درصد را در مطالعات مختلف نشان دادند. معماری ResNet که در ۸ مطالعه مورد استفاده قرار گرفته بود، به‌طور مکرر به‌عنوان یکی از بهترین مدل‌ها شناسایی شد. برتری معماری ResNet به دلیل استفاده از اتصالات میانبر

[Downloaded from jdm.tums.ac.ir on 2026-04-28]

(skip connections) است که مشکل محو شدن گرادیان در شبکه‌های عمیق را کاهش می‌دهد و امکان آموزش شبکه‌های عمیق‌تر را فراهم می‌کند. مدل‌های ترکیبی مانند ResNet+SAM (۱۶) و ResNet18+C (۱۲) که پارامترهای بالینی را با داده‌های تصویری ادغام می‌کردند، عملکرد حتی بهتری نشان دادند. در مدل ResNet+SAM، ماژول توجه فضایی (SAM) به مدل کمک می‌کند تا بر مناطق مهم‌تر تصویر تمرکز کند و ویژگی‌های مرتبط با پوسیدگی را با دقت بیشتری استخراج نماید. در مطالعه Ying و همکاران (۱۳)، معماری YOLOv5 به دلیل توازن مناسب بین سرعت و دقت تشخیص و توانایی شناسایی اشیاء در مقیاس‌های مختلف، عملکرد بهتری نسبت به سایر مدل‌ها داشت. این یافته‌ها اهمیت رویکردهای چند وجهی در تشخیص پوسیدگی دندان را برجسته می‌کند.

یکی دیگر از یافته‌های کلیدی، تأثیر قابل توجه تکنیک‌های پیش پردازش و افزایش داده بر عملکرد مدل‌ها بود. تقریباً تمام مطالعات از تغییر اندازه تصاویر و نرمال سازی استفاده کرده بودند و ۱۱ مطالعه روش‌های افزایش داده را برای بهبود عملکرد مدل و جلوگیری از بیش برازش به کار گرفته بودند. روش‌های نوآورانه‌ای مانند نمونه برداری مبتنی بر برش مرکز ثقل (۸) نیز برای بهبود دقت تشخیصی معرفی شده بودند. مقایسه بین الگوریتم‌های یادگیری ماشین و متخصصان انسانی، که در ۶ مطالعه انجام شده بود، نشان داد که این الگوریتم‌ها عملکرد مشابه یا حتی بهتری نسبت به متخصصان انسانی دارند. به طور خاص، در مطالعه Liu و همکاران (۱۶)، مدل ResNet+SAM عملکرد بهتری نسبت به دندانپزشکان تازه‌کار داشت، و در مطالعه Das و همکاران (۱۸)، نرم افزار مبتنی بر هوش مصنوعی دقت کلی بالاتری (۸۹ درصد در مقابل ۸۶ درصد) نسبت به تفسیر انسانی نشان داد. این یافته‌ها پتانسیل استفاده از ابزارهای مبتنی بر هوش مصنوعی به‌عنوان سیستم‌های پشتیبان تصمیم‌گیری در محیط‌های بالینی را نشان می‌دهد.

مزایا و محدودیت‌های استفاده از یادگیری ماشین در تشخیص پوسیدگی دندان

یادگیری عمیق مزایای قابل توجهی برای تشخیص پوسیدگی‌های دندانی ارائه می‌دهد. مهم‌ترین مزیت آن، دقت تشخیصی بالا حتی در شناسایی پوسیدگی‌های اولیه است که اغلب برای متخصصان انسانی چالش برانگیز است. برخلاف انسان‌ها که ممکن است تحت تأثیر خستگی، تجربه متفاوت یا سوگیری قرار گیرند، الگوریتم‌های یادگیری ماشین با ثبات و عینیت بیشتری عمل می‌کنند. سرعت بالای پردازش این الگوریتم‌ها – از چند میلی‌ثانیه تا چند ثانیه برای هر تصویر – آن‌ها را به ابزارهایی کارآمد برای کاربردهای بالینی تبدیل می‌کند.

استفاده از این فناوری می‌تواند بار کاری متخصصان را کاهش دهد و به آن‌ها اجازه دهد تمرکز بیشتری بر موارد پیچیده و تصمیم‌گیری‌های درمانی داشته باشند. همچنین، تشخیص زودهنگام پوسیدگی‌های اولیه امکان مداخلات پیشگیرانه را فراهم می‌کند و می‌تواند نیاز به درمان‌های تهاجمی‌تر را کاهش دهد. با وجود این مزایا، چالش‌ها و محدودیت‌هایی نیز وجود دارد. یکی از محدودیت‌های اصلی، نیاز به داده‌های آموزشی با کیفیت و حجم کافی است. شبکه‌های عصبی عمیق برای عملکرد مطلوب به هزاران تصویر با برچسب گذاری دقیق نیاز دارند. برچسب‌گذاری این داده‌ها، به ویژه برای پوسیدگی‌های اولیه، فرآیندی زمان بر و نیازمند تخصص بالاست. مطالعه Valenzuela و همکاران (۲۰) نشان داد که استفاده از میکرو–CT برای کمک به برچسب گذاری دقیق‌تر می‌تواند توافق بین‌ارزیاب و دقت تشخیصی را بهبود بخشد، اما این روش در محیط بالینی روزمره عملی نیست.

تعمیم پذیری مدل‌های آموزش دیده به شرایط واقعی نیز چالشی دیگر است. بیشتر مدل‌ها بر روی مجموعه داده‌های محدود و کنترل شده آموزش دیده‌اند و ممکن است در مواجهه با تنوع بیشتر تصاویر در محیط‌های بالینی واقعی عملکرد متفاوتی نشان دهند. همچنین، ماهیت "جعبه سیاه" الگوریتم‌های یادگیری عمیق، درک دقیق از نحوه تصمیم‌گیری آن‌ها را دشوار می‌سازد، که می‌تواند بر اعتماد متخصصان بالینی به این فناوری تأثیر بگذارد. نیاز به منابع محاسباتی قابل توجه برای آموزش و گاهی استفاده از مدل‌های یادگیری عمیق نیز می‌تواند چالش برانگیز باشد، هرچند در مطالعات بررسی‌شده، زمان تشخیص پس از آموزش مدل در حد میلی‌ثانیه تا ثانیه بوده است که برای کاربردهای بالینی کاملاً مناسب است.

چالش‌ها و پیشنهادات برای آینده

مرور مطالعات نشان می‌دهد که یادگیری عمیق در تشخیص پوسیدگی‌های دندانی با چالش‌های متعددی مواجه است که نیازمند توجه

[Downloaded from jdm.tums.ac.ir on 2026-04-28]

در پژوهش‌های آینده است. یکی از مهم‌ترین چالش‌ها، فقدان یک استاندارد مرجع قابل اعتماد و یکپارچه برای تشخیص پوسیدگی دندان است. استفاده از روش‌های پیشرفته مانند میکرو-CT که توسط Valenzuela و همکاران (۲۰) پیشنهاد شده، می‌تواند به ایجاد مجموعه داده‌های با کیفیت بالاتر کمک کند، اما دسترسی به این فناوری محدود است و کاربرد آن در محیط بالینی روزمره عملی نیست.

تنوع در مجموعه داده‌ها و معیارهای ارزیابی مورد استفاده در مطالعات مختلف، مقایسه مستقیم نتایج را دشوار می‌سازد. ایجاد مجموعه داده‌های استاندارد و در دسترس عموم با برچسب گذاری دقیق و قابل اعتماد می‌تواند به پیشرفت تحقیقات در این زمینه کمک کند. همچنین، توسعه چارچوب‌های استاندارد برای ارزیابی الگوریتم‌های تشخیص پوسیدگی دندان می‌تواند قابلیت مقایسه نتایج را بهبود بخشد.

علی‌رغم تلاش برای انجام یک مرور نظام‌مند جامع، این مطالعه با محدودیت‌هایی مواجه بود که باید در تفسیر نتایج مد نظر قرار گیرند. نخست، محدودیت زبانی این مرور که تنها شامل مقالات منتشر شده به زبان انگلیسی بود، می‌تواند منجر به از دست دادن برخی مطالعات مرتبط منتشر شده به سایر زبان‌ها شده باشد. دوم، علی‌رغم تلاش برای جستجوی جامع، ممکن است برخی مطالعات منتشر نشده یا مقالات کنفرانسی که در پایگاه‌های مورد جستجو نمایه نشده‌اند، از دسترس خارج مانده باشند. سوم، ناهمگنی در روش شناسی، مجموعه داده‌ها و معیارهای ارزیابی استفاده شده در مطالعات مختلف، مقایسه مستقیم نتایج را دشوار می‌سازد و مانع از انجام متاآنالیز می‌شود. چهارم، عدم دسترسی به داده‌های خام و کدهای مورد استفاده در برخی مطالعات، امکان باز تولید و تأیید یافته‌ها را محدود می‌کند. با وجود این محدودیت‌ها، تلاش شده است تا با ارزیابی دقیق کیفیت مطالعات و استخراج سیستماتیک داده‌ها، شواهد موجود در این زمینه به بهترین شکل ممکن ارائه شود.

پژوهش‌های آینده می‌توانند بر چند محور اصلی متمرکز شوند. نخست، انجام مطالعات با حجم نمونه بزرگتر و مطالعات چند مرکزی می‌تواند به افزایش قابلیت تعمیم الگوریتم‌ها در شرایط واقعی کمک کند. دوم، توسعه و ارزیابی روش‌های تلفیق داده‌های چندوجهی مانند ترکیب داده‌های رادیوگرافی با پارامترهای بالینی، که در مطالعات Zheng و همکاران (۱۲) و Ngnamsie و همکاران (۲) نتایج امید بخشی نشان داده است، می‌تواند دقت تشخیصی را بهبود بخشد. توسعه روش‌های بهتر برای تفسیر پذیری تصمیمات الگوریتم‌ها نیز از اهمیت ویژه‌ای برخوردار است. تکنیک‌هایی مانند Grad-CAM (Gradient-weighted Class Activation Mapping) که در برخی از مطالعات بررسی‌شده استفاده شده است، می‌توانند به متخصصان در درک بهتر نحوه تصمیم‌گیری الگوریتم‌ها کمک کنند و اعتماد آن‌ها را به این فناوری افزایش دهند. ارزیابی تأثیر کاربرد ابزارهای تشخیصی مبتنی بر یادگیری ماشین بر تصمیم‌گیری درمانی و نتایج بالینی نیز زمینه مهمی برای پژوهش‌های آینده است. مطالعاتی که به بررسی این موضوع می‌پردازند که آیا استفاده از این ابزارها منجر به تشخیص زود هنگام‌تر، کاهش درمان‌های تهاجمی غیرضروری و بهبود نتایج بالینی می‌شود یا خیر، می‌تواند به پذیرش گسترده‌تر این فناوری در محیط‌های بالینی کمک کند. نهایتاً، توجه به مسائل اخلاقی، قانونی و اجتماعی مرتبط با استفاده از هوش مصنوعی در تشخیص پزشکی و دندانپزشکی، از جمله مسائل مربوط به مسئولیت، حفظ حریم خصوصی و امنیت داده‌ها، نیز باید در پژوهش‌های آینده مورد توجه قرار گیرد.

یافته‌های این مرور نظام‌مند نشان می‌دهد که الگوریتم‌های یادگیری عمیق، به ویژه CNN، پتانسیل قابل توجهی برای بهبود تشخیص پوسیدگی‌های دندانی در رادیوگرافی‌های پری اپیکال و سایر تصاویر رادیوگرافی داخل دهانی دارند. این الگوریتم‌ها در اکثر مطالعات دقت بالایی (۸۲ درصد تا ۹۸ درصد) در تشخیص پوسیدگی دندان نشان داده‌اند و در برخی موارد عملکرد مشابه یا حتی بهتری نسبت به متخصصان انسانی داشته‌اند. معماری‌های ResNet و مدل‌های ترکیبی مبتنی بر آن (مانند ResNet+SAM و ResNet+C) در اکثر مطالعات بهترین عملکرد را نشان داده‌اند. همچنین، YOLOv5 در مطالعه Ying و همکاران (۱۳) با F1 score برابر با ۸۷٪ بهترین عملکرد را در میان معماری‌های مورد مقایسه داشته است. استفاده از تکنیک‌های پیش پردازش تصویر و افزایش داده نیز تأثیر قابل توجهی بر بهبود عملکرد مدل‌ها داشته است.

ترکیب داده‌های رادیوگرافی با پارامترهای بالینی، همان‌طور که در مطالعه Zheng و همکاران (۱۲) با مدل ResNet18+C و Ngnamsie و همکاران (۲) با مدل ترکیبی ANN و DenseNet-201 نشان داده شده است، می‌تواند منجر به بهبود قابل توجه دقت تشخیصی شود. همچنین، استفاده از روش‌های پیشرفته‌تر برای برچسب‌گذاری داده‌ها،

[Downloaded from jdm.tums.ac.ir on 2026-04-28]

مانند میکرو–CT در مطالعه Valenzuela و همکاران (۲۰)، می‌تواند به بهبود کیفیت داده‌های آموزشی و افزایش دقت مدل‌ها کمک کند.

با این حال، چالش‌هایی نیز در کاربرد یادگیری ماشین برای تشخیص پوسیدگی دندان وجود دارد. محدودیت در دسترس بودن داده‌های با کیفیت و برچسب گذاری شده، چالش‌های تعمیم پذیری به محیط‌های بالینی واقعی، و مسائل مربوط به تفسیر پذیری و اعتماد به الگوریتم‌ها از جمله این چالش‌ها هستند. همچنین، نحوه ادغام این ابزارها در جریان کار روزمره دندانپزشکان و مسائل اخلاقی و قانونی مربوط به استفاده از آن‌ها نیاز به بررسی بیشتر دارد.

استفاده از ابزارهای تشخیصی مبتنی بر یادگیری ماشین می‌تواند مزایای متعددی در محیط‌های بالینی داشته باشد. این ابزارها می‌توانند به تشخیص زود هنگام پوسیدگی‌های دندانی، به ویژه پوسیدگی‌های اولیه که ممکن است توسط متخصصان انسانی نادیده گرفته شوند، کمک کنند. همچنین، این ابزارها می‌توانند بار کاری متخصصان را کاهش داده و به آن‌ها اجازه دهند تا بر موارد پیچیده‌تر تمرکز کنند. استفاده از این ابزارها به عنوان یک سیستم پشتیبان تصمیم‌گیری، و نه جایگزینی برای متخصصان انسانی، می‌تواند به بهبود کیفیت مراقبت‌های دندانپزشکی کمک کند.

در مجموع، یافته‌های این مرور نظام‌مند نشان می‌دهد که یادگیری عمیق، به ویژه CNNها، ابزاری قدرتمند برای تشخیص پوسیدگی‌های دندانی در رادیوگرافی‌های پری اپیکال و سایر تصاویر رادیوگرافی داخل دهانی است. با پیشرفت فناوری و افزایش دسترسی به داده‌های با کیفیت، انتظار می‌رود که این ابزارها در آینده نقش مهم‌تری در تشخیص دندانپزشکی ایفا کنند. با این حال، پژوهش‌های بیشتری برای غلبه بر چالش‌های موجود و ارزیابی تأثیر استفاده از این ابزارها بر نتایج بالینی مورد نیاز است.

References:

1- Albano D, Galiano V, Basile M, Di Luca F, Gitto S, Messina C, et al. Artificial intelligence for radiographic imaging detection of caries lesions: a systematic review. BMC Oral Health. 2024;24(1):274.

2- Ngnamsie Njimbouom S, Lee K, Kim JD. MMDCP: Multi-modal dental caries prediction for decision support system using deep learning. Int J Environ Res Public Health. 2022;19(17):10928.

3- Szabó V, Szabo BT, Orhan K, Veres DS, Manulis D, Ezhov M, et al. Validation of artificial intelligence application for dental caries diagnosis on intraoral bitewing and periapical radiographs. J Dent. 2024;147:105105.

4- Van Leemput P, Keustermans J, Mollemans W. Statistical validation of a deep learning algorithm for dental anomaly detection in intraoral radiographs using paired data. arXiv. 2024;13(46):36.

5- Ayhan B, Ayan E, Atsü S. Detection of dental caries under fixed dental prostheses by analyzing digital panoramic radiographs with artificial intelligence algorithms based on deep learning methods. BMC Oral Health. 2025;25(1):216.

6- Schwendicke F, Tzschoppe M, Paris S. Radiographic caries detection: a systematic review and meta-analysis. J Dent. 2015;43(8):924-33.

7- Bennasar C, Nadal-Martínez A, Arroyo S, Gonzalez-Cid Y, López-González ÁA, Pérez F, et al. Integrating Machine Learning and Deep Learning for Predicting Non-Surgical Root Canal Treatment Outcomes Using Two-Dimensional Periapical Radiographs. Diagnostics (Basel). 2025;15(8):1009.

8- Qayyum A, Tahir A, Butt MA, Luke A, Abbas HT, Qadir J, et al. Dental caries detection using a semi-supervised learning approach. Sci Rep. 2023;13(1):749.

9- Asci E, Kilic M, Celik O, Cantekin K, Bircan HB, Bayrakdar İS, et al. A Deep Learning Approach to Automatic Tooth Caries Segmentation in Panoramic Radiographs of Children in Primary Dentition, Mixed Dentition, and Permanent Dentition. Children (Basel). 2024;11(6):690.

10- Lee JH, Kim DH, Jeong SN, Choi SH. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. J Dent. 2018;77:106-11.

11- Liu F, Gao L, Wan J, Lyu Z-L, Huang Y-Y, Liu C, et al. Recognition of Digital Dental X-ray Images Using a Convolutional Neural Network. J Digit Imaging. 2023;36:(1):73-9.

12- Zheng L, Wang H, Mei L, Chen Q, Zhang Y, Zhang H. Artificial intelligence in digital cariology: a new tool for the diagnosis of deep caries and pulpitis using convolutional neural networks. Ann Transl Mede. 2021;9(9):763.

13- Ying S, Huang F, Shen X, Liu W, He F. Performance comparison of multifarious deep networks on caries detection with tooth X-ray images. J Dent. 2024;144:104970.

14- Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.

15- Chen H, Li H, Zhao Y, Zhao J, Wang Y. Dental disease detection on periapical radiographs based on deep convolutional neural networks. Int J Comput Ass Radiol Surg. 2021;16(4):649-61.

16- Liu Y, Xia K, Cen Y, Ying S, Zhao Z. Artificial intelligence for caries detection: a novel diagnostic tool using deep learning algorithms. Oral Radiol. 2024;40(3):375-84.

17- Khan HA, Haider MA, Ansari HA, Ishaq H, Kiyani A, Sohail K, et al. Automated feature detection in dental periapical radiographs by using deep learning. Oral Surg Oral Med Oral Pathol Oral Radiol. 2021;131(6):711-20.

18- Das M, Shahnawaz K, Raghavendra K, Kavitha R, Nagareddy B, Murugesan S. Evaluating the accuracy of

[Downloaded from jdm.tums.ac.ir on 2026-04-28]

AI-based software vs human interpretation in the diagnosis of dental caries using intraoral radiographs: An RCT. J Pharm Bioallied Sci. 2024;16(Suppl 1):S812-S4.

19- Rampf S, Gehrig H, Möltner A, Fischer MR, Schwendicke F, Huth KC. Radiographical diagnostic competences of dental students using various feedback methods and integrating an artificial intelligence application-A randomized clinical trial. Eur J Dent Edu. 2024;28(4):925-37.

20- Valenzuela REG, Mettes P, Loos BG, Marquering H, Berkhout E. Enhancement of early proximal caries annotations in radiographs: introducing the Diagnostic Insights for Radiographic Early-caries with micro-CT (ACTA-DIRECT) dataset. BMC Oral Health. 2024;24(1):1325.

21- Cantu AG, Gehrung S, Krois J, Chaurasia A, Rossi JG, Gaudin R, et al. Detecting caries lesions of different radiographic extension on bitewings using deep learning. J Dent. 2020;100:103425.

22- Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529-36.

23- Powers DMW. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. Int J Machine Learning Tech. 2011;2(1):37-63.

[Downloaded from jdm.tums.ac.ir on 2026-04-28]