دانشنامه علمی چاوش طب آذین · chavoshteb.ir · به‌روزرسانی ۱ آذر ۱۴۰۵

تشخیص پوکی استخوان در رادیوگرافی پانورامیک با هوش مصنوعی: مرور سیستماتیک و متاآنالیز

هوش مصنوعی، به‌ویژه یادگیری عمیق، در تشخیص پوکی استخوان روی رادیوگرافی پانورامیک دقت بالایی نشان داده است: حساسیت تجمیعی ۸۷٫۹۲ درصد و ویژگی ۸۱٫۹۳ درصد در متاآنالیز ۲۴ مطالعه.[۱] با این حال، در مرور سیستماتیک DMFR هیچ مطالعه‌ای اعتبارسنجی خارجی یا پیاده‌سازی در مطب را گزارش نکرده بود؛[۲] بنابراین این ابزارها فعلاً کمک‌کننده ارجاع‌اند، نه جایگزین DXA.

نکات کلیدی این مقاله

- متاآنالیز Ghasemi و همکاران با ۲۴ مطالعه، حساسیت تجمیعی ۸۷٫۹۲ درصد، ویژگی ۸۱٫۹۳ درصد و نسبت شانس تشخیصی ۳۲٫۹۹ را گزارش کرد.[۱]
- در متاآنالیز Khadivi و همکاران (۱۲ مطالعه در تحلیل کمّی)، حساسیت تجمیعی ۰٫۸۵ و ویژگی ۰٫۹۵ بود.[۳]
- آزمون Egger شواهدی از اثر مطالعات کوچک و خطر سوگیری انتشار نشان داد (p کمتر از ۰٫۰۰۱).[۱]
- در مرور ۱۳ مطالعه یادگیری عمیق، DXA در ۸ مطالعه استاندارد مرجع بود و هیچ مطالعه‌ای اعتبارسنجی خارجی نداشت.[۲]
- نتیجه مدل در عمل باید به ارجاع برای DXA منجر شود، نه به تشخیص قطعی از روی پانورامیک.[۴]

پوکی استخوان با کاهش تراکم مواد معدنی استخوان و افزایش خطر شکستگی شناخته می‌شود و ریخت‌شناسی استخوان مندیبل در رادیوگرافی پانورامیک از مدت‌ها پیش برای غربالگری فرصت‌طلبانه آن به کار رفته است.[۱] شاخص‌های کیفی و کمّی کورتکس مندیبل در مقاله جداگانه دانشنامه بررسی شده‌اند؛ این مقاله فقط به مدل‌های هوش مصنوعی می‌پردازد و شواهد متاآنالیزها، نوع استاندارد مرجع و محدودیت‌های تعمیم‌پذیری را برای رادیولوژیست دهان و فک جمع‌بندی می‌کند.

فهرست مطالب

1. تشخیص پوکی استخوان با هوش مصنوعی: پرسش بالینی
2. نتایج متاآنالیزها
3. استاندارد مرجع و تعمیم‌پذیری

4. منابع سوگیری و محدودیت‌ها
5. نکات کاربردی برای رادیولوژیست
6. جدول جمع‌بندی نکات
7. پرسش‌های متداول
8. جمع‌بندی
9. مطالعه مرتبط در دانشنامه
10. منابع

تشخیص پوکی استخوان با هوش مصنوعی: پرسش بالینی

در این حوزه دو پرسش متفاوت مطرح است: آیا الگوریتم می‌تواند قضاوت یک رادیولوژیست باتجربه درباره فرسایش کورتکس مندیبل را بازتولید کند، و آیا می‌تواند وضعیت تراکم استخوان را در برابر DXA پیش‌بینی کند. پاسخ این دو پرسش یکسان نیست و تفسیر اعداد دقت بدون توجه به استاندارد مرجع گمراه‌کننده است.[۲، ۵]

در مطالعه Lee و همکاران، رادیولوژیست‌های دهان و فک با بیش از ۱۰ سال تجربه، پانورامیک ۱۲۶۸ زن را بررسی کردند و وجود فرسایش کورتکس تحتانی مندیبل را معیار پوکی استخوان گرفتند؛ سطح زیر منحنی سه مدل DCNN بین ۰٫۹۷۶۳ و ۰٫۹۹۹۱ بود.[۵] این عدد بیانگر هم‌خوانی با قضاوت رادیولوژیست است، نه دقت در برابر سنجش تراکم استخوان.

نتایج متاآنالیزها

متاآنالیز ۲۴ مطالعه روی پانورامیک

Ghasemi و همکاران مطالعات استفاده‌کننده از هوش مصنوعی برای تشخیص پوکی استخوان روی پانورامیک را تا اکتبر ۲۰۲۴ جست‌وجو و با QUADAS-2 و GRADE ارزیابی کردند. دقت مطالعات بین ۵۰ تا ۹۹ درصد، حساسیت بین ۵۰ تا ۱۰۰ درصد و ویژگی بین ۳۸ تا ۱۰۰ درصد متغیر بود؛ فقط ۱۰ مطالعه در همه حوزه‌ها خطر سوگیری پایین داشتند.[۱]

دانشنامه علمی چاوش طب آذین

دقت تجمیعی هوش مصنوعی در تشخیص پوکی استخوان از تصاویر دندانی

درصد؛ داده‌های دو متاآنالیز ۲۰۲۵



منبع: متاآنالیزهای Ghasemi و همکاران و Khadivi و همکاران (۲۰۲۵)

حساسیت و ویژگی تجمیعی در دو متاآنالیز ۲۰۲۵؛ نمودار تحریریه بر پایه داده‌های منتشرشده.

> «هوش مصنوعی، به‌ویژه یادگیری عمیق، دقت تشخیصی بالایی در تشخیص پوکی استخوان روی رادیوگرافی‌های پانورامیک نشان داد.»
>
> “AI, particularly deep learning, showed high diagnostic accuracy in detecting osteoporosis on panoramic radiographs.”
>
> — Ghasemi و همکاران، J Dent 2025[1]

متاآنالیز تصاویر دندانی و مقایسه الگوریتم‌ها

Khadivi و همکاران از ۶۴۰ رکورد، ۲۲ مطالعه را در مرور کیفی و ۱۲ مطالعه را در متاآنالیز وارد کردند. حساسیت تجمیعی ۰٫۸۵ (فاصله اطمینان ۹۵ درصد: ۰٫۷۰ تا ۰٫۹۳) و ویژگی تجمیعی ۰٫۹۵ (۰٫۹۱ تا ۰٫۹۷) بود. الگوریتم‌های متعارف حساسیت ۰٫۸۲ و ویژگی ۰٫۹۶ و شبکه‌های عصبی کانولوشنی عمیق حساسیت ۰٫۸۷ و ویژگی ۰٫۹۲ داشتند.[3]

نتایج دو متاآنالیز ۲۰۲۵ درباره هوش مصنوعی و پوکی استخوان

شاخص	Ghasemi و همکاران	Khadivi و همکاران
تصاویر واردشده	فقط پانورامیک	تصاویر دندانی

شاخص	Ghasemi و همکاران	Khadivi و همکاران
تعداد مطالعات	۲۴ مطالعه	۲۲ مطالعه کیفی؛ ۱۲ مطالعه در متاآنالیز
حساسیت تجمیعی	۸۷٫۹۲ درصد	۰٫۸۵ (۰٫۷۰ تا ۰٫۹۳)
ویژگی تجمیعی	۸۱٫۹۳ درصد	۰٫۹۵ (۰٫۹۱ تا ۰٫۹۷)
شاخص‌های دیگر	DOR ۳۲٫۹۹؛ LR+ ۴٫۸۷	CNN عمیق: حساسیت ۰٫۸۷، ویژگی ۰٫۹۲
هشدار روش‌شناختی	ناهمگنی زیاد؛ اثر مطالعات کوچک	نیاز به نمونه بزرگ‌تر و استانداردسازی تصویر

منبع: Ghasemi و همکاران، J Dent 2025؛ Khadivi و همکاران، Osteoporos Int 2025.

استاندارد مرجع و تعمیم‌پذیری

DXA یا قضاوت متخصص

در مرور سیستماتیک Querrer و همکاران در DMFR، از ۱۳ مطالعه واجد شرایط، DXA در ۸ مطالعه استاندارد مرجع بود و پانورامیک در ۱۲ مطالعه روش تصویربرداری اصلی بود. معماری‌های ResNet و EfficientNet، به‌ویژه همراه با روش‌های ترکیبی و متغیرهای بالینی، عملکرد بهتری داشتند.[۲]

Nakamoto و همکاران سه شبکه را با برچسب‌گذاری رادیولوژیست روی ۱۵۰۰ پانورامیک زنان ۵۰ ساله و بالاتر آموزش دادند و آن‌ها را روی ۱۰۰ بیمار دیگر با نتیجه سنجش تراکم استخوان ستون کمری و گردن فمور آزمودند. هم‌خوانی با قضاوت رادیولوژیست ۸۶ تا ۹۰٫۷ درصد بود، اما در برابر تراکم ستون کمری حساسیت ۷۸٫۳ تا ۸۲٫۶ درصد و ویژگی ۷۱٫۴ تا ۷۹٫۲ درصد به دست آمد.[۴] این فاصله نشان می‌دهد که دقت در برابر DXA معمولاً پایین‌تر از هم‌خوانی با برچسب انسانی است.

اثر استاندارد مرجع بر اعداد دقت در دو مطالعه پایه

مطالعه	استاندارد مرجع	نتیجه گزارش‌شده
Lee و همکاران ۲۰۱۹	فرسایش کورتکس مندیبل به تشخیص رادیولوژیست باتجربه	AUC ۰٫۹۷۶۳ تا ۰٫۹۹۹۱ در ۲۰۰ پانورامیک آزمون

مطالعه	استاندارد مرجع	نتیجه گزارش‌شده
Nakamoto و همکاران ۲۰۲۲ (برچسب انسانی)	طبقه‌بندی کورتکس توسط یک رادیولوژیست	هم‌خوانی ۸۶ تا ۹۰٫۷ درصد
Nakamoto و همکاران ۲۰۲۲ (ستون کمری)	سنجش تراکم استخوان	حساسیت ۷۸٫۳ تا ۸۲٫۶؛ ویژگی ۷۱٫۴ تا ۷۹٫۲ درصد
Nakamoto و همکاران ۲۰۲۲ (گردن فمور)	سنجش تراکم استخوان	حساسیت ۸۰ تا ۸۶٫۷؛ ویژگی ۶۷٫۱ تا ۷۴٫۱ درصد

منبع: Lee و همکاران، DMFR 2019؛ Nakamoto و همکاران، DMFR 2022.

> «با این حال، هیچ مطالعه‌ای اعتبارسنجی خارجی یا پیاده‌سازی در مطب دندانپزشکی را گزارش نکرد و این موضوع کاربردپذیری آن‌ها را محدود می‌کند.»
>
> “However, no study reported external validation or implementation in dental practice, limiting their applicability.”
>
> — Querrer و همکاران، Dentomaxillofac Radiol 2026[۲]

منابع سوگیری و محدودیت‌ها

در متاآنالیز Ghasemi، فراتحلیل رگرسیونی نشان داد اندازه نمونه فقط اثر حاشیه‌ای بر ناهمگنی دارد (R^2 برابر ۰٫۰۷۸) و آزمون Egger اثر مطالعات کوچک را مطرح کرد.[۱] به بیان دیگر، تفاوت در جمعیت، دستگاه، برچسب‌گذاری و طراحی مطالعه بخش بزرگی از پراکندگی نتایج را توضیح می‌دهد.

تطابق سنی گروه‌ها نیز اهمیت دارد: در مطالعه Lee میانگین سنی گروه بدون پوکی استخوان ۳۲٫۸ و گروه مبتلا ۷۲٫۲ سال بود،[۵] و چنین اختلافی می‌تواند به مدل اجازه دهد نشانه‌های سنی را به‌جای ویژگی بیماری بیاموزد؛ این برداشت تحریریه است و در خود مطالعه آزموده نشده است.

> «پژوهش‌های آینده باید اندازه نمونه در مجموعه داده‌های آموزش و آزمون را افزایش دهند و تکنیک‌های تصویربرداری را استاندارد کنند.»

> “Future research should expand sample sizes in test and training datasets and standardize imaging techniques”
>
> — Khadivi و همکاران، Osteoporos Int 2025[۳]

نکات کاربردی برای رادیولوژیست

- پیش از پذیرش هر عدد دقت، استاندارد مرجع را بررسی کنید: DXA، برچسب رادیولوژیست یا شاخص کورتیکال.[۲]
- خروجی مدل را نشانه‌ای برای ارجاع به سنجش تراکم استخوان در نظر بگیرید؛ مطالعات پایه نیز همین هدف را بیان کرده‌اند.[۴، ۵]
- عملکرد گزارش‌شده بدون اعتبارسنجی خارجی، به دستگاه و جمعیت دیگر قابل تعمیم فرض نشود.[۲]
- کیفیت پانورامیک و موقعیت‌دهی صحیح بیمار برای ارزیابی کورتکس مندیبل اهمیت دارد؛ خطاهای موقعیت‌دهی در مقاله مرتبط دانشنامه بررسی شده‌اند.

جدول جمع‌بندی نکات

نکات کلیدی هوش مصنوعی در تشخیص پوکی استخوان روی پانورامیک

محور	نکته کلیدی
دقت تجمیعی	حساسیت ۸۷٫۹۲٪ و ویژگی ۸۱٫۹۳٪ (۲۴ مطالعه)
متاآنالیز دوم	حساسیت ۰٫۸۵ و ویژگی ۰٫۹۵
استاندارد مرجع	DXA در ۸ از ۱۳ مطالعه یادگیری عمیق
اعتبارسنجی خارجی	در هیچ مطالعه‌ای گزارش نشده است
سوگیری	اثر مطالعات کوچک (آزمون Egger، p کمتر از ۰٫۰۰۱)
کاربرد امروز	ابزار کمکی ارجاع، نه جایگزین DXA

منبع: جمع‌بندی تحریریه بر پایه منابع فهرست‌شده در همین مقاله.

پرسش‌های متداول

آیا هوش مصنوعی می‌تواند پوکی استخوان را از روی پانورامیک تشخیص دهد؟

شواهد متاآنالیز دقت بالایی نشان می‌دهد، با حساسیت تجمیعی حدود ۸۸ درصد و ویژگی حدود ۸۲ درصد، اما مدل‌ها هنوز اعتبارسنجی خارجی ندارند.[۱، ۲]

حساسیت و ویژگی تجمیعی چقدر گزارش شده است؟

در یک متاآنالیز حساسیت ۸۷٫۹۲ و ویژگی ۸۱٫۹۳ درصد و در متاآنالیز دیگر حساسیت ۰٫۸۵ و ویژگی ۰٫۹۵ بود.[۱، ۳]

آیا شبکه‌های عصبی عمیق از الگوریتم‌های متعارف بهترند؟

در متاآنالیز Khadivi حساسیت CNN عمیق ۰٫۸۷ و الگوریتم متعارف ۰٫۸۲ بود و ویژگی‌ها ۰٫۹۲ و ۰٫۹۶ گزارش شد.[۳]

استاندارد مرجع در این مطالعات چیست؟

بیشتر مطالعات DXA را به کار برده‌اند، اما برخی قضاوت رادیولوژیست درباره کورتکس مندیبل را معیار گرفته‌اند.[۲، ۵]

چرا این مدل‌ها هنوز در مطب استفاده نمی‌شوند؟

هیچ مطالعه‌ای اعتبارسنجی خارجی یا پیاده‌سازی بالینی را گزارش نکرده و ناهمگنی بین مطالعات زیاد است.[۱، ۲]

نتیجه مثبت مدل چه اقدامی می‌طلبد؟

ارجاع بیمار برای سنجش تراکم استخوان، مانند DXA، و پیگیری پزشکی.[۴]

جمع‌بندی

هوش مصنوعی برای تشخیص پوکی استخوان روی رادیوگرافی پانورامیک در متاآنالیزها حساسیت و ویژگی امیدوارکننده‌ای دارد، اما بخش بزرگی از شواهد از مطالعات تک‌مرکزی با خطر سوگیری و بدون اعتبارسنجی خارجی است. ارزش واقعی این ابزارها در غربالگری فرصت‌طلبانه و ارجاع به موقع برای DXA است؛ رادیولوژیست دهان و فک باید پیش از اتکا به هر مدل، استاندارد مرجع، جمعیت آموزش و شیوه اعتبارسنجی آن را بررسی کند.

مطالعه مرتبط در دانشنامه

- شاخص کورتیکال مندیبل در پانورامیک و غربالگری پوکی استخوان
- هوش مصنوعی در رادیولوژی دندان و فک
- یافته‌های اتفاقی در رادیوگرافی پانورامیک
- بعد فرکتال و تحلیل استخوان ترابکولار
- خطاهای موقعیت‌دهی پانورامیک
- خوانش نقادانه مطالعات دقت تشخیصی

منابع

شماره‌های بالانویس در متن به همین فهرست پیوند دارند؛ هر منبع با شناسه DOI یا نشانی رسمی ناشر قابل بررسی است.

1. Ghasemi N, Rokhshad R, Zare Q, Shobeiri P, Schwendicke F. Artificial intelligence for osteoporosis detection on panoramic radiography: A systematic review and meta analysis. J Dent. 2025;156:105650. DOI PubMed

2. Querrer R, Vieira LS, Teodoro AB, Melo NS, Mesquita CRM, Silva MAG, et al. Deep learning for osteoporosis screening in dental practice: a systematic review. Dentomaxillofac Radiol. 2026;55(1):1-27. DOI PubMed

3. Khadivi G, Akhtari A, Sharifi F, Zargarian N, Esmaeili S, Ahsaie MG, et al. Diagnostic accuracy of artificial intelligence models in detecting osteoporosis using dental images: a systematic review and meta-analysis. Osteoporos Int. 2025;36(1):1-19. DOI PubMed

4. Nakamoto T, Taguchi A, Kakimoto N. Osteoporosis screening support system from panoramic radiographs using deep learning by convolutional neural network. Dentomaxillofac Radiol. 2022;51(6):20220135. DOI PubMed PMC

5. Lee JS, Adhikari S, Liu L, Jeong HG, Kim H, Yoon SJ. Osteoporosis detection in panoramic radiographs using a deep convolutional neural network-based computer-assisted diagnosis system: a preliminary study. Dentomaxillofac Radiol. 2019;48(1):20170344. DOI PubMed PMC

درباره نویسنده: تیم تحریریه علمی چاوش طب آذین

این مقاله را تیم تحریریه علمی چاوش طب آذین بر پایه منابع داوری‌شده و بیانیه‌های رسمی فهرست‌شده تهیه کرده است؛ ادعای بازبینی بالینی مستقل ندارد و جایگزین قضاوت متخصص نیست. آخرین به‌روزرسانی: ۱ آذر ۱۴۰۵.