دانشنامه علمی چاوش طب آذین · chavoshteb.ir · به‌روزرسانی ۲۸ آبان ۱۴۰۵

تشخیص خودکار لندمارک‌های سفالومتری با یادگیری عمیق: دقت دوبعدی و سه‌بعدی

لندمارک‌های سفالومتری در فراتحلیل ۱۹ مطالعه یادگیری عمیق، به نسبت ۰٫۷۹۹ در آستانه ۲ میلی‌متری از مرجع انسانی تعیین شدند و فراتحلیل دیگری با ۴۰ مطالعه توافق ۷۹ درصد در ۲ میلی‌متر و ۹۰ درصد در ۳ میلی‌متر گزارش کرد. مرور چتری ۱۱ مرور سیستماتیک نتیجه گرفت که خطای بیشتر لندمارک‌ها از حد توصیه‌شده بالاتر است و نظارت متخصص لازم است.

نکات کلیدی این مقاله

- در فراتحلیل Schwendicke، نسبت لندمارک‌های تعیین‌شده در آستانه ۲ میلی‌متر ۰٫۷۹۹ بود و خطر سوگیری بیشتر مطالعات بالا بود.[۱]
- نرخ موفقیت تجمعی در آستانه‌های ۱، ۲، ۲٫۵، ۳ و ۴ میلی‌متر به ترتیب ۶۵، ۸۱، ۸۶، ۹۱ و ۹۶ درصد گزارش شد.[۶]
- از ۴۰ مطالعه فراتحلیل Mesquita فقط سه مطالعه در همه حوزه‌ها خطر سوگیری پایین داشتند.[۵]
- مرور چتری نشان داد تعیین خودکار لندمارک در CBCT کم‌دقت‌تر از رادیوگرافی دوبعدی است.[۲]
- یک مدل سه‌بعدی با خطای شعاعی میانگین زیر ۱٫۳ میلی‌متر، مهارت متخصصان ارشد و تازه‌کار را ۱۵٫۹ و ۲۸٫۹ درصد بهبود داد.[۴]

تعیین دستی لندمارک‌ها زمان‌بر و وابسته به تجربه مشاهده‌گر است و یادگیری عمیق برای خودکارسازی آن روی سفالوگرام دوبعدی و CBCT به کار رفته است.[۱] پرسش بالینی این است که خطای الگوریتم در چه حدی است و آیا معیار رایج ۲ میلی‌متری برای تصمیم درمانی کافی است.[۲] این مقاله شواهد فراتحلیل‌ها و مطالعات سه‌بعدی را برای رادیولوژیست دهان و فک جمع‌بندی می‌کند.

فهرست مطالب

1. لندمارک‌های سفالومتری و معیارهای سنجش خطا
2. شواهد فراتحلیل‌ها در سفالوگرام دوبعدی

3. تعیین لندمارک سه‌بعدی در CBCT و CT
4. محدودیت آستانه ۲ میلی‌متر
5. همکاری انسان و هوش مصنوعی
6. نکات کاربردی برای رادیولوژیست
7. جدول جمع‌بندی نکات
8. پرسش‌های متداول
9. جمع‌بندی
10. مطالعه مرتبط در دانشنامه
11. منابع

لندمارک‌های سفالومتری و معیارهای سنجش خطا

دقت الگوریتم‌های تعیین لندمارک‌های سفالومتری معمولاً با دو شاخص گزارش می‌شود: خطای شعاعی میانگین (MRE)، یعنی فاصله اقلیدسی نقطه پیش‌بینی‌شده از نقطه مرجع، و نرخ موفقیت تشخیص (SDR)، یعنی درصد لندمارک‌هایی که در شعاع مشخصی مانند ۲ میلی‌متر قرار می‌گیرند.[۳]، [۴] مرجع در بیشتر مطالعات علامت‌گذاری یک یا دو متخصص است و خود این مرجع خطای بین‌مشاهده‌گر دارد.[۱]



سفالوگرام جانبی؛ لندمارک‌های استخوانی، دندانی و بافت نرم روی چنین تصویری به‌صورت دستی یا خودکار تعیین می‌شوند و خطای چند میلی‌متری می‌تواند اندازه‌گیری‌ها را تغییر دهد. تصویر: Challiyan، Wikimedia Commons؛ مجوز CC BY-SA 4.0، با برش و تغییر اندازه.

شواهد فراتحلیل‌ها در سفالوگرام دوبعدی

فراتحلیل Schwendicke

Schwendicke و همکاران ۱۹ مطالعه منتشرشده بین ۲۰۱۷ و ۲۰۲۰ را تحلیل کردند که همگی شبکه عصبی کانولوشنی به کار برده بودند؛ ۱۵ مطالعه روی رادیوگرافی جانبی دوبعدی و ۱۲ مطالعه با داده‌های عمومی بودند و به‌طور میانگین ۳۰ لندمارک بررسی شد. خطای پیش‌بینی حول آستانه ۲ میلی‌متری متمرکز بود و نسبت لندمارک‌های داخل این آستانه ۰٫۷۹۹ (۰٫۷۷۰ تا ۰٫۸۲۴) به دست آمد.[۱]

> «مجموعه شواهد سازگار است، اما از خطر سوگیری بالا رنج می‌برد.»
>
> “The overall body of evidence is consistent but suffers from high risk of bias.”
>
> — Schwendicke و همکاران، Clin Oral Investig 2021[۱]

فراتحلیل‌های Mesquita و Londono

در فراتحلیل Mesquita و همکاران با ۴۰ مطالعه، توافق هوش مصنوعی با علامت‌گذاری دستی ۷۹ درصد در آستانه ۲ میلی‌متر و ۹۰ درصد در آستانه ۳ میلی‌متر بود؛ ناهمگونی بسیار بالا (I^2 برابر ۹۹ درصد) و قطعیت شواهد بسیار پایین گزارش شد و لندمارک منتون کمترین اختلاف را داشت.[۵] Londono و همکاران با ۲۱ مطالعه، نرخ موفقیت تجمعی را در آستانه‌های ۱ تا ۴ میلی‌متر از ۶۵ تا ۹۶ درصد گزارش کردند.[۶]

دانشنامه علمی چاوش طب آذین

نرخ موفقیت تشخیص لندمارک‌های سفالومتری با یادگیری ماشین

درصد لندمارک‌های داخل هر آستانه خطا در سفالوگرام دوبعدی (فراتحلیل ۲۱ مطالعه)



منبع: Londono و همکاران، Saudi Dent J 2023.

نرخ موفقیت تشخیص لندمارک در آستانه‌های مختلف؛ نمودار تحریریه بر پایه فراتحلیل Londono (۲۰۲۳).

فراتحلیل‌های تشخیص خودکار لندمارک‌های سفالومتری

مطالعه	تعداد مطالعات	نتیجه اصلی	محدودیت
Schwendicke ۲۰۲۱	۱۹	نسبت ۰٫۷۹۹ در ۲ میلی‌متر	سوگیری بالا، داده سه‌بعدی اندک
Mesquita ۲۰۲۳	۴۰	۷۹٪ در ۲ و ۹۰٪ در ۳ میلی‌متر	قطعیت بسیار پایین
Londono ۲۰۲۳	۲۱	۸۱٪ در ۲ میلی‌متر	دقت متغیر بین مطالعات
Polizzi ۲۰۲۴، مرور چتری	۱۱ مرور	خطای بیشتر لندمارک‌ها بالاتر از حد توصیه	آستانه ۲ میلی‌متر نادرست

منبع: مطالعات فهرست‌شده در همین مقاله.

تعیین لندمارک سه‌بعدی در CBCT و CT

Schwendicke در ۲۰۲۱ داده‌های سه‌بعدی را اندک اما امیدوارکننده توصیف کرد.[۱] Liu و همکاران یک مدل U-Net سه‌بعدی سبک را روی ۴۸۰ CT مارپیچی و ۲۴۰ CBCT آموزش دادند و روی ۳۲۰ CT و ۱۵۰ CBCT دیگر آزمودند؛ MRE در هر دو روش زیر ۱٫۳ میلی‌متر و در شرایط دشوار مانند مال‌اکلوژن، فقدان دندان و آرتیفکت فلزی زیر ۱٫۴ میلی‌متر بود. در CBCT لندمارک‌های دندانی دقیق‌تر از لندمارک‌های استخوانی بودند و بیشترین خطا در محور کرونال دیده شد.[۴]

با این حال مرور چتری Polizzi و Leonardi نتیجه گرفت که تعیین خودکار لندمارک روی CBCT کم‌دقت‌تر از رادیوگرافی دوبعدی است؛ بنابراین نتایج یک مطالعه چندمرکزی را باید با احتیاط به همه سامانه‌ها تعمیم داد.[۲]

محدودیت آستانه ۲ میلی‌متر

Polizzi و Leonardi با ارزیابی ۱۱ مرور سیستماتیک با ابزار ROBIS نشان دادند که هوش مصنوعی همه لندمارک‌ها را با دقت یکسان تعیین نمی‌کند و نتیجه‌گیری بیشتر مطالعات بر آستانه نادرست ۲ میلی‌متری استوار بوده است.[۲] برداشت تحریریه این است که گزارش SDR در ۲ میلی‌متر بدون تفکیک لندمارک، خطای لندمارک‌های دشوارتر را که روی زوایا و فواصل سفالومتری اثر می‌گذارند پنهان می‌کند.[۲، ۵]

> «تا امروز، تعیین لندمارک سفالومتری مبتنی بر هوش مصنوعی همچنان به نظارت نهایی یک ارتودنتیست باتجربه نیاز دارد.»
>
> “To date, AI-driven cephalometric landmarking still requires the final supervision of an experienced orthodontist.”
>
> — Polizzi و Leonardi، J Dent 2024[۲]

همکاری انسان و هوش مصنوعی

Le و همکاران مدلی را روی ۱۱۹۳ سفالوگرام برای ۴۱ لندمارک آموزش دادند که MRE برابر ۱٫۸۷ میلی‌متر و SDR برابر ۷۳٫۱۷ درصد در ۲ میلی‌متر داشت. همکاری دانشجویان مبتدی با مدل، SDR را ۵٫۳۳ درصد و نرخ طبقه‌بندی صحیح را ۸٫۳۸ درصد افزایش داد.[۳] در مطالعه Liu نیز مدل سه‌بعدی زمان تعامل با رابط کاربری را ۶ تا ۹٫۵ برابر کوتاه کرد.[۴]

شاخص‌های عملکرد در دو مطالعه اصلی

مطالعه	داده	MRE	اثر بر کاربر
Le ۲۰۲۲	۱۱۹۳ سفالوگرام، ۴۱ لندمارک	۱٫۸۷ میلی‌متر	افزایش SDR مبتدی‌ها ۵٫۳۳٪
Liu ۲۰۲۵	CT و CBCT چندمرکزی	زیر ۱٫۳ میلی‌متر	بهبود ۱۵٫۹٪ و ۲۸٫۹٪ مهارت متخصصان

منبع: Le و همکاران 2022؛ Liu و همکاران 2025.

نکات کاربردی برای رادیولوژیست

برداشت تحریریه از این شواهد:[۱، ۵، ۲، ۴]

- خروجی خودکار را پیش‌نویس بدانید و لندمارک‌های پرخطا را پیش از گزارش اندازه‌گیری‌ها بازبینی کنید.[۲]
- هنگام انتخاب نرم‌افزار، SDR تفکیک‌شده برای هر لندمارک و نتایج آزمون خارجی را درخواست کنید، نه فقط میانگین کلی.[۱، ۵]
- در CBCT، خطای محور کرونال و لندمارک‌های استخوانی را جداگانه بررسی کنید.[۴]
- کیفیت سفالوگرام و موقعیت‌دهی بیمار همچنان پیش‌شرط هر تحلیل خودکار است.[۱]

جدول جمع‌بندی نکات

نکات کاربردی تشخیص خودکار لندمارک‌های سفالومتری

محور	نکته کلیدی
دقت دوبعدی	حدود ۸۰ درصد در آستانه ۲ میلی‌متر
آستانه ۳ میلی‌متر	حدود ۹۰ درصد
کیفیت شواهد	سوگیری بالا و ناهمگونی شدید
CBCT	دقت کمتر از دوبعدی در مرور چتری

محور	نکته کلیدی
معیار ۲ میلی‌متر	برای تصمیم بالینی ناکافی
نقش متخصص	بازبینی نهایی ضروری

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

پرسش‌های متداول

دقت یادگیری عمیق در تعیین لندمارک‌های سفالومتری چقدر است؟

حدود ۸۰ درصد لندمارک‌ها در فاصله ۲ میلی‌متری از مرجع انسانی قرار می‌گیرند.[۱، ۶]

SDR و MRE چه تفاوتی دارند؟

MRE میانگین فاصله نقطه پیش‌بینی‌شده از مرجع است و SDR درصد لندمارک‌های داخل یک شعاع مشخص.[۳]

آیا آستانه ۲ میلی‌متر کافی است؟

مرور چتری این آستانه را برای نتیجه‌گیری بالینی نادرست دانست، زیرا خطای لندمارک‌ها یکسان نیست.[۲]

دقت تعیین خودکار لندمارک در CBCT چگونه است؟

یک مدل چندمرکزی MRE زیر ۱٫۳ میلی‌متر داشت، اما مرور چتری دقت CBCT را کمتر از دوبعدی گزارش کرد.[۲، ۴]

آیا هوش مصنوعی می‌تواند جایگزین متخصص شود؟

خیر؛ نظارت نهایی متخصص باتجربه هنوز لازم است.[۲]

همکاری انسان و هوش مصنوعی چه سودی دارد؟

در دانشجویان مبتدی SDR را ۵٫۳۳ درصد افزایش داد و در مطالعه سه‌بعدی مهارت متخصصان را بهبود بخشید.[۳، ۴]

جمع‌بندی

تشخیص خودکار لندمارک‌های سفالومتری با یادگیری عمیق در سفالوگرام دوبعدی به دقتی حدود ۸۰ درصد در آستانه ۲ میلی‌متر رسیده و مدل‌های سه‌بعدی جدید نتایج امیدوارکننده‌ای دارند. با این حال سوگیری بالای مطالعات، ناهمگونی شدید و ناکافی بودن معیار ۲ میلی‌متری نشان می‌دهد که خروجی الگوریتم باید زیر نظر متخصص بازبینی شود و ارزیابی هر لندمارک جداگانه انجام گیرد.

مطالعه مرتبط در دانشنامه

- سفالومتری لترال: تکنیک و پایایی لندمارک‌ها
- هوش مصنوعی در رادیولوژی دندان و فک
- بلوغ مهره‌های گردنی (CVM) در سفالومتری
- سگمنتیشن دندان در CBCT با هوش مصنوعی
- خوانش نقادانه مطالعات دقت تشخیصی
- CBCT در ارتودنسی: شواهد توجیه

منابع

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

1. Schwendicke F, Chaurasia A, Arsiwala L, Lee JH, Elhennawy K, Jost-Brinkmann PG, et al. Deep learning for cephalometric landmark detection: systematic review and meta-analysis. Clin Oral Investig. 2021;25(7):4299-4309. DOI PubMed PMC

2. Polizzi A, Leonardi R. Automatic cephalometric landmark identification with artificial intelligence: An umbrella review of systematic reviews. J Dent. 2024;146:105056. DOI PubMed

3. Le VNT, Kang J, Oh IS, Kim JG, Yang YM, Lee DW. Effectiveness of Human-Artificial Intelligence Collaboration in Cephalometric Landmark Detection. J Pers Med. 2022;12(3):387. DOI PubMed PMC

4. Liu B, Liu C, Xiong Y, Zhu H, Zeng W, Chen J, et al. Accuracy and reliability of 3D cephalometric landmark detection with deep learning. Eur J Med Res. 2025;30(1):1000. DOI PubMed PMC

5. de Queiroz Tavares Borges Mesquita G, Vieira WA, Vidigal MTC, Travençolo BAN, Beaini TL, Spin-Neto R, et al. Artificial Intelligence for Detecting Cephalometric Landmarks: A Systematic Review and Meta-analysis. J Digit Imaging. 2023;36(3):1158-1179. DOI PubMed PMC

6. Londono J, Ghasemi S, Hussain Shah A, Fahimipour A, Ghadimi N, Hashemi S, et al. Evaluation of deep learning and convolutional neural network algorithms accuracy for detecting and predicting anatomical landmarks on 2D lateral cephalometric images: A systematic review and meta-analysis. Saudi Dent J. 2023;35(5):487-497. 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: ۲۸ آبان ۱۴۰۵.