

تخمین سن استخوانی کودکان مبتنی بر یادگیری عمیق با استفاده از رادیوگرافی دست

مجتبی سیرتی-امشه^۱، الهام شعبانی‌نیا^۲، علی چاپاریان^۳

مقاله پژوهشی

چکیده

مقدمه: از تصاویر رادیوگرافی دست، به صورت رایج جهت ارزیابی بلوغ استخوانی استفاده می‌شود. به طوری که تفاوت چشمگیر میان سن ارزیابی شده و سن تقویمی می‌تواند نشان‌دهنده اختلال در رشد باشد. با این حال ارزیابی دستی تصاویر، معمولاً فرایندی زمان‌بر و وابسته به ناظر است. لذا هدف از انجام این مطالعه، ایجاد روشی خودکار برای ارزیابی سن استخوانی با استفاده از تصاویر رادیوگرافی دست می‌باشد.

روش‌ها: در این پژوهش که از نوع بنیادی- کاربردی می‌باشد از مجموعه تصاویر رادیوگرافی انجمن رادیولوژی آمریکای شمالی (Radiological Society of North America) RSNA استفاده شد و روش یادگیری انتقالی برای تخمین سن استخوانی کودکان پیشنهاد گردید. تصاویر ورودی، ابتدا به دلیل کیفیت پایین مورد پیش‌پردازش قرار گرفتند. سپس مدل از پیش آموزش دیده DenseNet-۱۲۱ برای استخراج ویژگی‌های مکانی متمایزکننده مورد استفاده قرار گرفت.

یافته‌ها: ارزیابی‌ها با استفاده از پنج مدل از پیش آموزش دیده و بر روی مجموعه‌ی داده‌ی RSNA نشان دادند که مدل DenseNet-۱۲۱ پس از تنظیم می‌تواند با میانگین خطای مطلق ۹/۸ ماه بهتر از سایر مدل‌ها عمل کند.

نتیجه‌گیری: بلوغ اسکلتی می‌تواند با استفاده از مدل DenseNet-۱۲۱ با دقت رضایت‌بخشی تخمین زده شود و از این روش می‌توان به رادیولوژیست‌ها در اندازه‌گیری سریع و دقیق سن استخوانی کمک نمود.

واژگان کلیدی: اختلالات رشد؛ اندازه‌گیری؛ یادگیری عمیق؛ سن استخوانی؛ رادیوگرافی

ارجاع: سیرتی-امشه مجتبی، شعبانی‌نیا الهام، چاپاریان علی. تخمین سن استخوانی کودکان مبتنی بر یادگیری عمیق با استفاده از رادیوگرافی

دست. مجله دانشکده پزشکی اصفهان ۱۴۰۱؛ ۴۰ (۷۰۰): ۱۰۴۳-۱۰۳۷

مقدمه

و دارای خطا باشد. علاوه بر این، نظرات افراد مختلف در شرایط گوناگون می‌تواند متغیر باشد (تغییرات درون ناظر و بین ناظر [Intraobserver and interobserver]). در نتیجه، یک روش سریع و خودکار برای تخمین سن استخوانی می‌تواند بسیار مفید واقع گردد. در گذشته، تخمین خودکار سن استخوانی از طریق رویکردهای سنتی و از طریق استفاده از روش‌های یادگیری ماشین مانند ماشین بردار پشتیبان (Support vector machine) (۱)، k نزدیک‌ترین همسایه (K-Nearest Neighbor) (۲) مرسوم بوده است. متأسفانه رویکردهای سنتی، سطح دقت پایینی ارائه می‌کنند و در بیشتر موارد به دانش قبلی برای استخراج ویژگی‌های مناسب از تصاویر دست نیاز

برای بررسی ناهنجاری‌های رشد در کودکان که به دلایل متعدد از جمله اختلالات ژنتیکی یا غدد درون‌ریز ممکن است ایجاد شوند، از روش ارزیابی سن استخوانی استفاده می‌شود (۱). ارزیابی بلوغ استخوانی با استفاده از تصاویر رادیوگرافی دست، به دلیل سادگی و نیاز به دوز تابشی متوسط، یک روش معمول می‌باشد. دو روش رایج برای این کار GP (Greulich and Pyle) (۲) و TW (Tanner-Whitehouse) (۳) هستند. رویکرد GP، رایج‌ترین روشی است که عموماً توسط متخصصین رادیولوژی استفاده می‌شود که بر مقایسه‌ی رادیوگرافی با یک اطلس مرجع تمرکز دارد (۱). استفاده از هر یک از این روش‌ها می‌تواند زمان‌بر

۱- دانشجوی کارشناسی ارشد، گروه فیزیک پزشکی، دانشکده‌ی پزشکی، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران

۲- استادیار، گروه ریاضی کاربردی، دانشکده‌ی علوم و فناوری‌های نوین، دانشگاه تحصیلات تکمیلی و فناوری پیشرفته، کرمان، ایران

۳- استاد، گروه فیزیک پزشکی، دانشکده‌ی علوم پزشکی اصفهان، اصفهان، ایران

نویسنده‌ی مسؤو: علی چاپاریان: استاد، گروه فیزیک پزشکی، دانشکده‌ی پزشکی، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران

Email: ali_chaparian@yahoo.com

معماری‌ها روی مجموعه داده‌ای بسیار کوچک‌تر امکان‌پذیر می‌شود. روش یادگیری انتقالی به ویژه در این مطالعه می‌تواند مناسب باشد زیرا تصاویر رادیوگرافی دست، اغلب نسبت به بسیاری از تصاویر دیگر مانند تصاویر موجود در مجموعه‌ی داده‌ی ImageNet (۱۳) پیچیدگی کمتری دارند. بنابراین در این مطالعه، اثر انتقال یادگیری بر روی مجموعه‌ی داده‌ی RSNA مورد بررسی قرار گرفت.

پیش‌پردازش: از آن‌جا که تصاویر موجود در مجموعه‌ی داده‌ی RSNA از چندین بیمارستان با استفاده از دستگاه‌های مختلف تصویربرداری تهیه شده بود، مشکلاتی مانند وضوح مختلف، توزیع ناهموار سطوح خاکستری رنگ‌ها، نویز و غیره می‌توانست دقت مراحل بعدی را با مشکل روبرو نماید. معمولاً تصاویر ورودی ظاهری رنگ پریده، با محدوده‌ی باریکی از سطوح خاکستری بودند که جزئیات بافت استخوانی دست در آن چندان نمایان نبود. لذا برای رسیدگی به این مشکلات، پیش‌پردازش‌هایی بر روی این مجموعه‌ی داده انجام گردید که ترتیب اعمال این فیلترها از مطالعه‌ی Gao و همکاران (۱۰) الهام گرفته شد:

۱- یکسان‌سازی هیستوگرام (Histogram equalization): در این مرحله، توزیع کلی سطوح خاکستری تصویر یکنواخت شده و جزئیات بهبود یافت.

۲- تیز کردن یا واضح‌سازی تصویر (Sharpening): در این مرحله، از فیلتر لاپلاسی برای یافتن لبه‌ها، کانتورها و جزئیات تصویر استفاده شد. این جزئیات مجدداً با تصویر ورودی جمع شدند تا تصویر واضحی از تصویر ورودی ایجاد شود. استفاده از فیلتر لاپلاسی در عین حال که تغییرات شدید سطوح خاکستری را به خوبی آشکار می‌سازد و می‌تواند سبب تقویت نویز نیز بشود. از این‌رو قبل از اعمال لاپلاسی، فیلتر گاوسین (Gaussian) روی تصویر اعمال شد تا نویز تصویر ورودی تا حد امکان کاهش یابد.

۳- تصحیح گاما (Gamma correction): اگرچه استفاده از یکسان‌سازی هیستوگرام می‌تواند کنتراست تصویر را بهبود داده و سطوح خاکستری را گسترش دهد، اما این تبدیل اساساً یک تبدیل خطی است. لذا برای تمرکز بر نواحی استخوانی نیاز بود تا از یک تبدیل غیرخطی، که بتواند جزئیات این نواحی را آشکار نماید، استفاده شود. این تبدیل به شکل زیر عمل می‌کند:

$$O = I^\gamma \quad (1)$$

که در آن I شدت روشنایی پیکسل ورودی، O شدت روشنایی پیکسل خروجی و γ یک عدد ثابت است. در کلیه‌ی آزمایشات مقدار گاما بزرگتر از ۱ و به شکل تجربی برابر با ۲/۵ انتخاب شد، تا نواحی مختلف مربوط به بافت استخوان که معمولاً روشن‌تر از زمینه هستند،

دارند. در سال‌های اخیر، روش‌های یادگیری عمیق، کاربرد گسترده‌ای در زمینه‌ی پزشکی، به ویژه در رادیولوژی داشته‌اند (۶). شبکه‌های عصبی کانولوشنی قادر هستند تا به طور خودکار سلسله‌مراتب ویژگی‌های متمایز را یاد بگیرند. از این‌رو تخمین سن استخوانی مبتنی بر روش‌های یادگیری عمیق به دلیل توانایی در درک خودکار ویژگی‌های متمایز رادیوگرافی مورد توجه قرار گرفته است (۷، ۸).

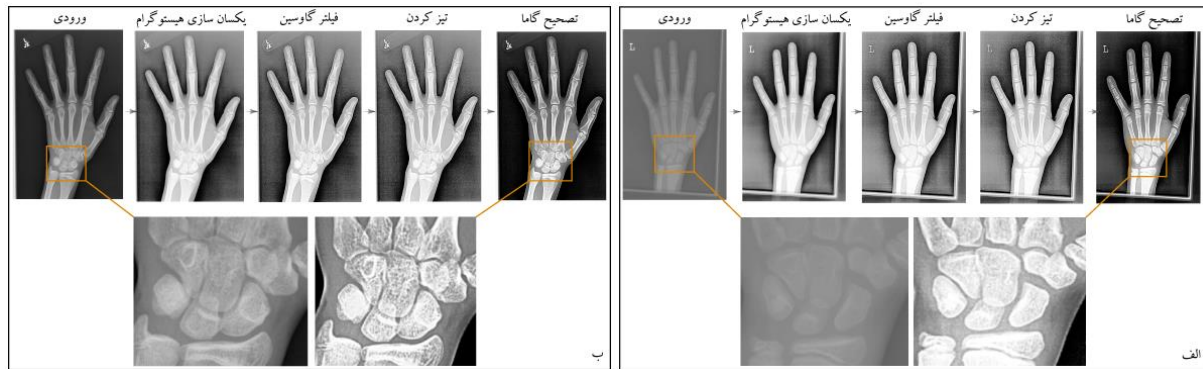
BoNet (۱) یکی از اولین مطالعاتی بود که یک شبکه‌ی عصبی کانولوشنی برای ارزیابی سن استخوانی مبتنی بر یادگیری عمیق ایجاد کرد. Wibisono و همکاران (۹) و Gao و همکاران (۱۰)، از مدل‌های شبکه‌ی عصبی کانولوشنی از پیش آموزش دیده برای روش یادگیری عمیق استفاده کردند. با توجه به اینکه روش‌های مبتنی بر یادگیری عمیق توانسته‌اند نتایج چشمگیری نسبت به روش‌های سنتی در حوزه‌های مختلف پزشکی به دست آورند، در این پژوهش، استفاده از روش یادگیری انتقالی برای تخمین سن استخوانی مورد توجه قرار گرفته است. هدف از این مطالعه، ایجاد روشی برای ارزیابی خودکار سن استخوان با استفاده از رادیوگرافی دست کودکان می‌باشد.

روش‌ها

مجموعه‌ی داده: در این مطالعه که از نوع بنیادی-کاربردی می‌باشد، از مجموعه تصاویر رادیوگرافی انجمن رادیولوژی آمریکای شمالی RSNA (Radiological Society of North America) (۱۱) که برای عموم در دسترس است، استفاده گردید. این تصاویر توسط تعدادی رادیولوژیست با تجربه و با استفاده از روش GP برچسب گذاری شد. ۱۲۶۱۱ تصویر در مجموعه‌ی داده‌ی RSNA با سنین ۱ تا ۲۲۸ ماه وجود دارد که این مجموعه داده به صورت تصادفی به ۸۱۹۷، ۴۲۱۴ و ۲۰۰ تصویر به ترتیب برای مجموعه‌ی آموزش، مجموعه‌ی اعتبارسنجی و مجموعه‌ی تست تقسیم شدند.

تمامی آزمایش‌ها بر روی رایانه‌ای شامل یک پردازنده‌ی گرافیکی NVIDIA RTX ۲۰۶۰ با ۶ گیگابایت رم انجام شد. کدهای اجرا شده در پایتون با استفاده از کتابخانه Keras 28.0 نوشته شده و از TensorFlow به عنوان بک‌اند (Backend) استفاده گردید.

یادگیری انتقالی: یادگیری انتقالی، روشی است که در آن، یک مدل توسعه یافته برای یک مسئله، به عنوان نقطه‌ی شروع برای یک مدل در یک مسئله دوم دوباره استفاده می‌شود (۱۲). معماری‌های مختلفی از شبکه‌های عصبی کانولوشنی وجود دارند که با موفقیت بر روی مجموعه‌ی داده‌های بزرگ آموزش داده شده‌اند و به آن‌ها اجازه می‌دهند تا ویژگی‌هایی مانند لبه‌ها، کانتورها (Contours) و رنگ‌ها را به طور مؤثر از تعداد زیادی تصویر استخراج کنند. بنابراین توانایی استخراج اطلاعات مرتبط و ارائه‌ی پیش‌بینی‌های دقیق، با تنظیم این



شکل ۱. شکل‌های (الف) و (ب) نشان‌دهنده دو نمونه تصویر بهبود یافته بعد از پیش پردازش‌های انجام شده بر روی تصویر ورودی

$$MAE = \frac{1}{N} \sum_{i=1}^N |f_i - y_i| \quad (2)$$

$$MSE = \frac{1}{N} \sum_{i=1}^N (f_i - y_i)^2 \quad (3)$$

که N تعداد کل نمونه‌ی داده تست را نشان می‌دهد، f_i سن واقعی استخوان و y_i سن پیش‌بینی شده می‌باشد.

یافته‌ها

پارامترهای مورد استفاده در آموزش مدل DenseNet-۱۲۱ تنظیم شده، در جدول ۱ نشان داده شده است. بر این اساس، بهینه‌ساز Adam (Adam) برای آموزش مدل ما با نرخ یادگیری اولیه ۰/۰۰۱، اندازه‌ی دسته ۱۶ و مقدار وزن تنظیم L_2 برابر با ۰/۰۰۰۱ انتخاب شد.

جدول ۱. پارامترهای مدل از پیش آموزش دیده

پارامتر	مقدار
تابع فعال‌سازی (Activation function) خروجی	خطی (Linear)
بهینه‌ساز (Optimizer)	آدام (Adam)
تابع اتلاف (Loss function)	خطای مربعات میانگین
معیار ارزیابی (Metric)	خطای مطلق میانگین
اندازه‌ی دسته (Batch size)	۱۶
نرخ یادگیری اولیه (Learning rate)	۰/۰۰۱
وزن رگولاریزاسیون L_2	۰/۰۰۰۱
اندازه‌ی تصویر	۲۸۰×۲۸۰

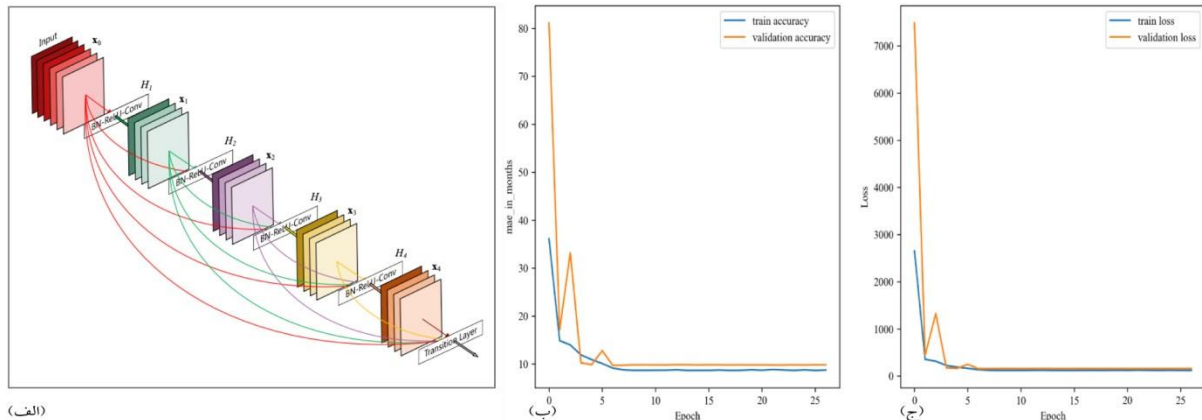
در شکل ۲ (ب و ج)، منحنی‌های خطای مطلق میانگین و خطای مربعات میانگین مدل نهایی، در طول آموزش و اعتبارسنجی مشاهده می‌شود. با توجه به نتایج نشان داده شده در این شکل، هر دو منحنی، کاهش رضایت‌بخشی را تجربه می‌کنند و در نهایت به حالت پایدار می‌رسند.

کاملاً از یکدیگر مجزا شده و در عین حال نواحی مربوط به زمینه کاملاً تیره شود. شکل ۱، دو نمونه از تصاویر ورودی به همراه خروجی فیلترهای اعمال شده در هر مرحله و در نهایت تصاویر بهبود یافته را نمایش می‌دهد.

معماری ۱۲۱-DenseNet در این مطالعه، مدل از پیش‌آموزش داده شده DenseNet-۱۲۱ (۱۴) برای مسأله‌ی تخمین سن استخوانی مورد استفاده قرار گرفت. مدل DenseNet-۱۲۱ ترکیبی از لایه‌های متراکم (Dense) است که در هر لایه‌ی متراکم یک لایه‌ی کانولوشنی (Convolutional) و یک لایه‌ی ادغام (Pooling) وجود دارد. هر لایه در یک معماری پیشرو (Feed-forward) توسط لایه‌های کانولوشنی تماماً متصل به لایه‌های دیگر متصل می‌شود (۱۴). بنابراین هر لایه، ورودی‌هایی را از تمام لایه‌های قبلی دریافت می‌کند و نقشه‌های ویژگی خود را به تمام لایه‌های بعدی ارسال می‌نماید. مدل DenseNet-۱۲۱ مزایای مختلفی از جمله توانایی حل مشکل ناپدید شدن گرادینان و توانایی استفاده مجدد از ویژگی‌ها برای ایجاد شبکه‌های متراکم‌تر با پارامترهای کمتر را ارائه می‌دهد. همانطور که در شکل ۲ (الف) نشان داده شده است، در معماری DenseNet، هر لایه به طور مستقیم با هر لایه‌ی دیگر متصل است.

پس از فریز کردن (Freezing) لایه‌های مدل DenseNet-۱۲۱ از پیش آموزش داده شده، لایه‌ی ادغام میانگین‌گیری سراسری (Global average pooling) برای جلوگیری از برازش بیش از حد (Overfitting) و کاهش تعداد پارامترهای مدل انتخاب گردید و لایه‌های متراکم بعد از لایه‌ی ادغام میانگین قرار گرفتند. سپس از رگولاریزاسیون (Regularization) L_2 برای کاهش برازش بیش از حد و اجتناب از یادگیری مدل‌های پیچیده استفاده شد (۱۵).

معیارهای ارزیابی: خطای مطلق میانگین و خطای مربعات میانگین معیارهایی هستند که به منظور ارزیابی عملکرد مدل از پیش آموزش دیده DenseNet-۱۲۱ در نظر گرفته شدند:



شکل ۲. نشان‌دهنده‌ی (الف) شماتیک معماری DenseNet، (ب) منحنی‌های خطای مطلق میانگین برای مجموعه‌ی آموزش و مجموعه‌ی اعتبارسنجی و (ج) منحنی‌های خطای مربعات میانگین را برای مجموعه‌ی آموزشی و مجموعه‌ی اعتبارسنجی

دقیق برای هدف تخمین سن استخوانی تنظیم گردید، از سایر مدل‌های شبکه‌ی عصبی کانولوشنی از پیش آموزش دیده شناخته شده، که مورد ارزیابی قرار گرفتند، عملکرد بهتری داشت و همچنین، تصاویر بهبود یافته نتیجه مطلوبی نسبت به تصاویر اصلی به دست آوردند.

نتایج این مطالعه نشان داد که یادگیری انتقالی می‌تواند به دلیل سرعت بالا و عدم نیاز به حجم زیاد داده، روشی سودمند برای یادگیری داده‌های استخوانی دست به شمار آید. همچنین یافته‌های تجربی نشان می‌دهند، مدل DenseNet احتمالاً به دلیل داشتن اتصالات متراکم می‌تواند ویژگی‌های متمایزکننده بهتری نسبت به سایر مدل‌ها به دست آورد. اثر این الگوی اتصال متراکم این است که به پارامترهای کمتری نسبت به شبکه‌های کانولوشنی سنتی نیاز دارد، زیرا نیازی به یادگیری مجدد نقشه‌های ویژگی اضافی نیست. معماری DenseNet پیشنهادی، به صراحت بین اطلاعاتی که به شبکه اضافه می‌شود و اطلاعاتی که حفظ می‌شوند، تفاوت قائل می‌شود. علاوه بر کارایی بهتر پارامترها، یکی دیگر از مزیت‌های بزرگ مدل‌های DenseNet این است که به دلیل جریان اطلاعات و گرادینان‌های بهبود یافته در سراسر شبکه، آموزش آن‌ها آسان‌تر است (۱۴).

علاوه بر این، روش ارائه شده در این مطالعه با برخی از شبکه‌های عصبی کانولوشنی آموزش دیده مانند ResNet-50 (۱۶)، VGG-16 (۱۷)، Xception (۱۸) و InceptionV3 (۱۹) بر روی مجموعه‌ی داده ImageNet با استفاده از رویکرد یادگیری انتقالی مقایسه شد. نتایج هم برای تصاویر اصلی و هم برای تصاویر بهبود یافته با اندازه‌های 256×256 و 280×280 در جدول ۲ نشان داده شده است.

بحث

در این مطالعه، به بررسی کاربرد یادگیری انتقالی و مدل‌های از پیش آموزش دیده بر روی تصاویر رادیوگرافی دست برای ارزیابی سن استخوانی پرداخته شد. یک مدل از پیش آموزش دیده DenseNet-121 برای تخمین خودکار سن استخوانی مورد استفاده قرار گرفت و یافته‌ها در مجموعه‌ی داده‌ی عمومی RSNA رضایت‌بخش بود. در مرحله‌ی اول، تصاویر اصلی و تصاویر بهبود یافته به منظور بررسی اثر کیفیت تصاویر در مجموعه‌ی داده به طور جداگانه به DenseNet-121 داده شد. همانطور که در جدول ۲ مشاهده شد، مدل از پیش آموزش دیده DenseNet-121 که به طور

جدول ۲. مقایسه‌ی نتایج شبکه‌های عصبی کانولوشنی از پیش آموزش دیده بر روی مجموعه داده RSNA برای اندازه‌های تصویر 256×256 و 280×280

تصاویر بهبود یافته (280×280)		تصاویر اصلی (280×280)		تصاویر بهبود یافته (256×256)		تصاویر اصلی (256×256)		مدل از پیش آموزش دیده
MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	
۲۳/۹۰	۹۶۴/۰۵	۲۸/۸۱	۱۳۹۰/۹۶	۲۲/۵	۸۶۹/۲	۲۸/۶	۱۳۳۲/۴	ResNet-50 (۱۶)
۱۶/۵۷	۴۶۰/۹۸	۱۷/۹۲	۵۱۵/۷۲	۱۸/۴	۵۶۳/۷	۱۹/۱	۶۶۱/۲	VGG-16 (۱۷)
۱۴/۷۰	۳۶۸/۴۷	۱۵/۵۶	۳۸۹/۹۳	۱۵/۵	۴۳۴/۱	۱۶/۹	۴۹۸/۸	Xception (۱۸)
۱۵/۴۴	۳۸۶/۱۹	۱۵/۹۸	۴۲۱/۵۷	۱۶/۲	۴۷۲/۵	۱۶/۱	۴۱۶/۳	InceptionV-3 (۱۹)
۹/۸	۱۶۵/۱۸	۱۰/۲۱	۱۸۵/۲۶	۱۰/۲	۱۷۱/۸	۱۰/۵	۱۷۹/۸	DenseNet-121 (۱۴)

مجموعه‌ی داده‌های بیشتری مورد بررسی قرار گیرد.

نتیجه‌گیری

با توجه به یافته‌های ارائه شده، نشان داده شد که یادگیری انتقالی می‌تواند تخمین سن استخوانی را به طور رضایت‌بخشی انجام دهد و همچنین بهبود تصاویر رادیوگرافی دست و پیش پردازش‌های انجام شده، موجب بهتر شدن تخمین سن استخوانی گردید. بنابراین، رویکرد پیشنهادی این مطالعه با استفاده از مدل DenseNet-۱۲۱ تنظیم شده، می‌تواند در اختلالات احتمالی که بر رشد و تکامل نوجوانان تأثیر می‌گذارد، به رادیولوژیست در تشخیص سریع و دقیق سن استخوانی کمک نماید.

تشکر و قدردانی

مطالعه‌ی حاضر، حاصل پایان‌نامه‌ی کارشناسی ارشد با شماره طرح تحقیقاتی ۳۴۰۱۰۴ می‌باشد که در حوزه‌ی معاونت پژوهشی دانشگاه علوم پزشکی اصفهان به تصویب رسید. بدین وسیله، نویسندگان این مقاله از حمایت‌های آن معاونت تشکر و قدردانی می‌کنند.

روش یادگیری انتقالی ارائه شده در این مطالعه به خطای مطلق میانگین ۹/۸ ماه دست یافت و این روش با دو مطالعه‌ی مرتبط Wibisono و همکاران (۹) و Gao و همکاران (۱۰) مقایسه شد، که هر دو از مجموعه‌ی داده RSNA استفاده کردند. Gao و همکاران (۱۰)، به خطای مطلق میانگین ۹/۹۹۷ ماه دست یافتند که نتیجه‌ی آن‌ها نزدیک به روش ارائه شده در مطالعه‌ی حاضر بود، ولی برای تقسیم‌بندی ناحیه‌ی استخوان دست، U-Net (۲۰) را انتخاب کردند و VGG-۱۶ از پیش آموزش داده شده را به عنوان شبکه‌ی پایه (Backbone) مدل به کار گرفتند. Wibisono و همکاران (۹) از مدل‌های شبکه‌ی عصبی کانولوشنی از پیش آموزش دیده VGG-۱۶ و MobileNets (۲۱) برای روش یادگیری عمیق استفاده کردند و به خطای مطلق میانگین ۱۴/۷۸ ماه رسیدند.

مقایسه‌ی روش‌های مختلف تخمین سن استخوان در کارهای مرتبط با مطالعه‌ی حاضر دشوار بود. زیرا از مجموعه داده‌ها و روش‌های ارزیابی متفاوتی استفاده کرده بودند و انجام یک مقایسه‌ی منصفانه، چالش‌برانگیز بود. لذا پیشنهاد می‌گردد در مطالعات آینده،

References

- Spampinato C, Palazzo S, Giordano D, Aldinucci M, Leonardi R. Deep learning for automated skeletal bone age assessment in X-ray images. *Med Image Anal* 2017; 36: 41-51.
- Greulich WW, Pyle SI. Radiographic atlas of skeletal development of the hand and wrist. 1st ed. California, UD: Stanford University Press; 1999.
- Poznanski AK. Assessment of skeletal maturity and prediction of adult height. *Am J Dis Child* 1977; 131(9): 1041-2.
- Harmsen M, Fischer B, Schramm H, Seidl T, Deserno TM. Support vector machine classification based on correlation prototypes applied to bone age assessment. *IEEE Journal of Biomedical and Health Informatics* 2012; 17(1): 190-7.
- Fischer B, Welter P, Grouls C, Günther RW, Deserno TM. Bone age assessment by content-based image retrieval and case-based reasoning. *Int J Comput Assist Radiol Surg* 2012; 7(3): 389-99.
- Fritz B, Marbach G, Civardi F, Fucentese SF, Pfirrmann CW. Deep convolutional neural network-based detection of meniscus tears: comparison with radiologists and surgery as standard of reference. *Skeletal Radiol* 2020; 49(8): 1207-17.
- He J, Jiang D. Fully automatic model based on se-resnet for bone age assessment. *IEEE Access*. 2021; 9: 62460-6.
- Zulkifley MA, Mohamed NA, Abdani SR, Kamari NAM, Moubark AM, Ibrahim AA. Intelligent bone age assessment: an automated system to detect a bone growth problem using convolutional neural networks with attention mechanism. *Diagnostics (Basel)* 2021; 11(5): 765.
- Wibisono A, Saputri MS, Mursanto P, Rachmad J, Yudasubrata ATW, Rizki F, et al. Deep learning and classic machine learning approach for automatic bone age assessment. *Proceedings of the 4th Asia-Pacific Conference on Intelligent Robot Systems (ACIRS)*; 2019: IEEE; Nagoya, Japan.
- Gao Y, Zhu T, Xu X. Bone age assessment based on deep convolution neural network incorporated with segmentation. *Int J Comput Assist Radiol Surg* 2020; 15(12): 1951-62.
- Halabi SS, Prevedello LM, Kalpathy-Cramer J, Mamontov AB, Bilbily A, Cicero M, et al. The RSNA pediatric bone age machine learning challenge. *Radiology* 2019; 290(2): 498-503.
- Tsiakmaki M, Kostopoulos G, Kotsiantis S, Ragos O. Transfer learning from deep neural networks for predicting student performance. *Appl Sci* 2020; 10(6): 2145.
- Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. Imagenet: A large-scale hierarchical image database. *Proceedings of the IEEE conference on computer vision and pattern recognition*; 2009 Jun 20-25; Miami, FL, USA.
- Huang G, Liu Z, van der Maaten L, Weinberger KQ. Densely connected convolutional networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*; 2017. Available from: URL: <https://arxiv.org/abs/1608.06993v5>
- van Laarhoven T. L2 regularization versus batch and weight normalization. 2017. Available from: URL: <https://arxiv.org/abs/1706.05350>
- He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. 2016. Available from: URL:

- <https://arxiv.org/abs/1512.03385>
17. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. 2014. Available from: URL: <https://arxiv.org/abs/1409.1556>
 18. Chollet F. Deep learning with depthwise separable convolutions. 2017. Available from: URL: <https://arxiv.org/abs/1610.02357>
 19. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016 Jun 27-30; Las Vegas, NV, USA; 2016.
 20. Ronneberger O, Fischer P, Brox T, Olaf R. U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Navab N, Hornegger J, Wells WM, Frangi AF, Editors. New York, NY: Springer International Publishing; 2015.
 21. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications. 2017. Available from: URL: <https://arxiv.org/abs/1704.04861>

Deep Learning-Based Pediatric Bone Age Estimation Using Hand Radiography

Mojtaba Sirati-Amsheh¹, Elham Shabaninia², Ali Chaparian³

Original Article

Abstract

Background: Hand radiographs are commonly used to evaluate bone maturity. So that the significant difference between the estimated bone age and the chronological age can indicate a developmental disorder. However, the manual evaluation of images is usually a time-consuming and observer-dependent process. Therefore, in this paper, an automatic method for the assessment of bone age using radiographs of children's hands is proposed.

Methods: In this fundamental-applied research, the collection of radiographic images of the Radiological Society of North America (RSNA) was used, and transfer learning methods were proposed. The input images were first pre-processed due to low quality. Then a pre-trained model based on DenseNet-121 was used to extract the discriminating spatial features.

Findings: Evaluations using five pre-trained models on the RSNA dataset showed that the DenseNet-121 model, after adjustment, could perform better than other models, with a mean absolute error of 9.8 months.

Conclusion: Skeletal maturity can be estimated with satisfactory accuracy using the DenseNet-121 model, and this method can help radiologists in quick and accurate measurement of bone age.

Keywords: Deep learning; Growth disorders; Bone age; Measurements; Radiography

Citation: Sirati-Amsheh M, Shabaninia E, Chaparian A. **Deep Learning-Based Pediatric Bone Age Estimation Using Hand Radiography.** J Isfahan Med Sch 2023; 40(700): 1037-43.

1- MSc Student of Medical Physics, Department of Medical Physics, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran

2- Assistant Professor, Department of Applied Mathematics, School of Sciences and Modern Technologies, Graduate University of Advanced Technology, Kerman, Iran

3- Professor, Department of Medical Physics, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran

Corresponding Author: Ali Chaparian, Professor, Department of Medical Physics, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran; Email: ali_chaparian@yahoo.com