

مروری بر انواع شبکه‌های عصبی به کار گرفته شده در بررسی تصاویر رادیوگرافی سراسرنما: مروری روایتی

چکیده

دریافت: ۱۴۰۳/۰۷/۱۸ ویرایش: ۱۴۰۳/۰۷/۲۵ پذیرش: ۱۴۰۳/۰۸/۲۳ آنلاین: ۱۴۰۳/۰۹/۰۱

با گسترش هوش مصنوعی، از انواع شبکه‌های عصبی مصنوعی به‌عنوان ابزاری برای ایجاد قدرت تحلیل در سیستم‌های کامپیوتری استفاده شد. مطالعاتی که با این هدف در حوزه دهان و دندان انجام شده‌اند، عموماً با تجزیه و تحلیل داده‌های حجیم سروکار دارند. در این مطالعه روش‌های مورد استفاده در تحلیل هوشمند و خودکار داده‌های دندانپزشکی، مورد بررسی قرار گرفته است. مطالعات بررسی شده به صورت مشترک از تصاویر رادیوگرافی سراسرنما (Ortho Pantoma Gram, OPG) استفاده و با سه هدف کلی تشخیص دندان (تفکیک هر دندان از دندان‌های قبل و بعد)، بخش‌بندی نمونه‌ای (تجزیه اطلاعات بصری تصویر به صورت قطعه قطعه) و بخش‌بندی معنایی (تحلیل و فهم اطلاعات استخراج شده از تصویر) انجام شده‌اند. برای حصول این موارد، وابسته به هدف پژوهش از مدل‌های مختلف شبکه‌های عصبی مصنوعی استفاده شده است. در این مطالعه، به مقایسه عملکرد شبکه‌های عصبی مصنوعی به کار گرفته شده، پرداخته شده است. به‌طور کلی تصاویر ورودی، ابتدا جهت آماده‌سازی دادگان برای خوانش توسط سیستم، پیش‌پردازش و سپس به شبکه عصبی تعلیم داده شده‌اند. در برخی موارد خروجی شبکه عصبی، به جهت بهبود نتایج، مجدد وارد چرخه پردازش شده است. در مقالات مورد بررسی به منظور تحلیل داده‌های دندانپزشکی، به‌طور کلی دو رویکرد دنبال شده است. در رویکرد حل مسئله به روش هوشمند، بیشتر از شبکه‌های عصبی پیچشی و در رویکرد خودکار، از روش‌های کلاسیک و بدون یادگیری استفاده شده است، در مطالعاتی که بر مبنای شبکه‌های عصبی مصنوعی انجام شده است سیستم با درک تصویر، بخش‌بندی، استخراج ویژگی، طبقه‌بندی ویژگی‌ها و مدل‌سازی شبکه و تنظیم متغیرهای آن، به یادگیری آنچه در هدف مطالعه تعیین شده پرداخته است.

کلمات کلیدی: شبکه‌های عصبی مصنوعی، تشخیص دهان و دندان، رادیوگرافی پانورامیکس.

مهدیه سلطانی^{۱*}، سیده زهره
سیدصالحی^{۲*}، ریحانه مهدوی^۱

۱- گروه مهندسی پزشکی، دانشکده پیراپزشکی،
دانشگاه علوم پزشکی آزاد تهران، تهران، ایران.

۲- گروه مهندسی پزشکی، دانشکده فنی
مهندسی، دانشگاه شاهد، تهران، ایران.

* نویسنده مسئول: تهران، دانشگاه علوم پزشکی آزاد
تهران، دانشکده پیراپزشکی، گروه مهندسی پزشکی.
تلفن: ۰۲۱-۲۲۰۰۶۶۶۰
E-mail: z.seyyedsalehi@iautmu.ac.ir

روزافزون استفاده از منابع هوشمند و طرح‌ریزی فرایندهای شناختی انسانی در رایانه‌ها، تحولی عظیم در دیدگاه محققان صنایع مختلف از جمله پزشکی ایجاد کرد. یادگیری ماشین (Machine Learning, ML) یک مجموعه روش برای القای یادگیری به سیستم است. در واقع این روش امکان آموزش را بدون اینکه به برنامه‌نویسی صریحی برای فهم سیستم نیاز باشد، فراهم می‌سازد. تمرکز اصلی یادگیری ماشینی بر توسعه برنامه‌های رایانه‌ای است که بتوانند به داده‌ها دسترسی پیدا

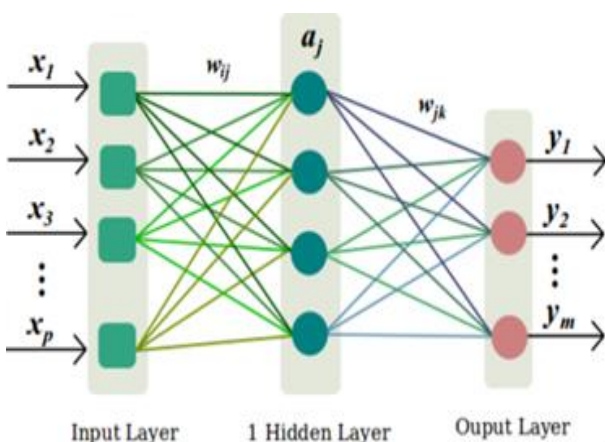
با ایجاد منابع هوشمند و خودکار با سازوکار شبکه‌های عصبی مصنوعی، در جهان و صنایع مختلف، کم‌کم این ایده مطرح شد که با استفاده از تحلیل خودکار داده‌های پزشکی، می‌توان به روند تشخیص و در مراحل پیشرفته‌تر به درمان بیماری‌ها، سرعت و دقت بیشتری بخشید. شبکه‌های عصبی مصنوعی با الهام از نحوه یادگیری مغز انسان، اطلاعات را تحلیل و با استفاده از استخراج ویژگی مشترک داده‌ها، به یادگیری آنچه باید تشخیص دهند، می‌پردازند. گسترش

(Integral Projection) برای بررسی خودکار تصاویر OPG استفاده شده است.^{۱۱}

مطالعات در حوزه روش‌های هوشمند با سه هدف کلی تشخیص دندان (تفکیک هر دندان از دندان قبل و بعد)، بخش‌بندی نمونه (تجزیه اطلاعات بصری تصویر بصورت قطعه به قطعه) و بخش‌بندی معنایی (تحلیل و ایجاد تمایز بین اطلاعات استخراج‌شده از تصویر) بررسی شده است در تشخیص سنتی نیز هر یک از این سه، یک مرحله از روندی است که درمانگر برای برنامه‌ریزی درمان انجام می‌دهد.^{۱۲}

از آنجایی که در حوزه پزشکی، بررسی دقیق‌تر ویژگی‌های موجود در داده‌ها ارزشمندتر از دیگر حوزه‌ها به نظر می‌رسد، اطلاعات تصاویر پزشکی به‌صورت مرحله به مرحله و هر بار با در نظر گرفتن یک یا چند ویژگی متفاوت بررسی می‌شود. به همین علت در تعداد قابل‌توجهی از مطالعات بررسی‌شده از الگوهایی با جزئی‌نگری بیشتر استفاده شده است. شبکه‌های عصبی عمیق نوعی از شبکه‌های قابل‌یادگیری هستند که تعداد لایه‌های آنها به‌منظور بررسی دقیق‌تر و ابعاد بیشتر، زیادتر شده است و نتیجه بررسی در هر لایه به عنوان ورودی لایه بعد، به منظور رسیدن به ویژگی‌های سطح بالا تعریف می‌شود.^{۱۳، ۱۴} شکل ۱ نمونه‌ای از یک شبکه عصبی با یک لایه پنهان را نشان می‌دهد.

از طرفی در اغلب پژوهش‌های پیشین استفاده از شبکه‌های عصبی پیچشی اولویت دارد.



کنند و از آن برای یادگیری خود استفاده کند. هدف اساسی الگوریتم‌های یادگیری ماشین، تعمیم یادگیری‌ها به فراتر از نمونه‌های آموزش داده شده است و این نشان‌دهنده‌ی مسیر درست تفسیر داده‌هاست.^۲

همچنین مورد دیگری که سبب افزایش ظرفیت‌های موجود در این فضا شده، استفاده از شبکه‌های عصبی عمیق (Deep neural networks, DNNs) است. این نوع شبکه‌ها با تعداد لایه‌های تحلیلی بیشتر^۳، به دلیل فراهم کردن امکان پیاده‌سازی فرآیندهای شناختی پیچیده و تحت تاثیر شبیه‌سازی تفکر جزئی‌تر همانند آنچه انسان در تحلیل مسائل پیچیده‌تر انجام می‌دهد، به تحلیل دقیق‌تر اطلاعات ضروری می‌پردازد.^۴ به همین دلیل این مدل‌ها، یکی از پرکاربردترین انواع شبکه در تحلیل داده‌های پزشکی به‌شمار می‌آید.^۵

شبکه عصبی پیچشی یکی از انواع پرکاربرد شبکه‌های عصبی مصنوعی عمیق در تحلیل و تشخیص جزئیات تصاویر پزشکی به‌ویژه تصاویر رادیوگرافی است. این مدل که به‌طور گسترده در پردازش تصویر و گفتار به‌کار می‌رود، در حوزه تصویر با بهره‌گیری از بینایی کامپیوتر و یادگیری ماشین امکان فرایند پردازش تصاویر را همانند یک تحریک دیداری شبیه‌سازی می‌کند.^۶

در این مرور تلاش می‌شود مطالعات انجام شده در حوزه‌ی شبیه‌سازی هوشمند عملکرد دندانپزشک یا درمانگر در مواجهه با تصاویر تشخیصی و درمانی، با رویکرد بررسی روش‌های مورد استفاده در تحلیل هوشمند تصاویر رادیوگرافی سراسرنما (Ortho Pantoma Gram, OPG) مورد بررسی قرار گیرد.^۷

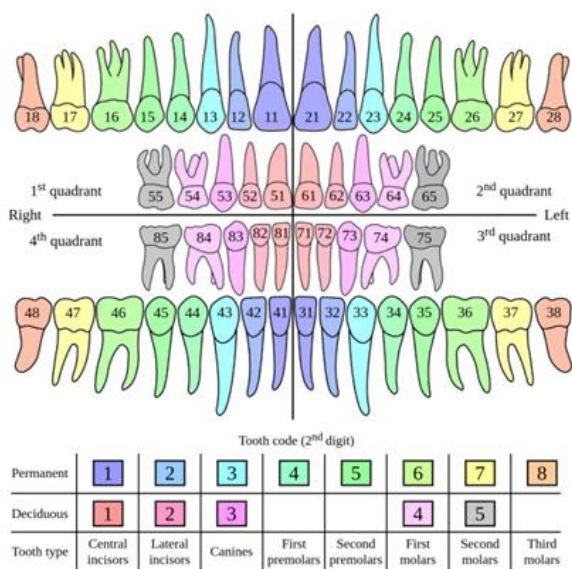
فهرست اصطلاحات تخصصی و اختصارات علمی شامل:

(Artificial Intelligence, AI)، (Ortho Pantoma Gram, OPG) و

(Convolutional Neural Network, CNN) می‌باشد.

همان‌طور که پیش‌تر نیز اشاره شد در این مرور، مطالعاتی با محوریت بررسی هوشمند و خودکار تصاویر OPG دهان و دندان، بررسی شده است. مقالات مورد بررسی، تحلیل کامپیوتری را با دو روش خودکار و هوشمند انجام داده‌اند.^{۸-۱۰} در روش خودکار در واقع سیستم بدون آگاهی و تجزیه‌وتحلیل تصاویر، صرفاً براساس آنچه به‌عنوان معیار تشخیص برای آن مشخص شده است، کار می‌کند، برای مثال در محوریت استخراج ویژگی‌های آناتومیک فک پایین، از روش آشکارساز لبه (Canny) و بازسازی برش‌نگاری

از مرزبندی المان‌های شناسایی شده که شامل دندان‌ها یا مرزهای فک می‌شود، شبکه این امکان را دارد که با ایجاد تمایز بین هر دو دندان، به شماره‌گذاری دندان‌ها بپردازد. سبک شماره‌گذاری به دلخواه انتخاب و به شبکه اعمال می‌شود تا خروجی تعیین شود. یکی از این سبک‌های شماره‌گذاری براساس سیستم FDI می‌باشد (شکل ۳).^{۱۸،۱۹}



شکل ۳: شماره گذاری دندان‌ها در سیستم FDI

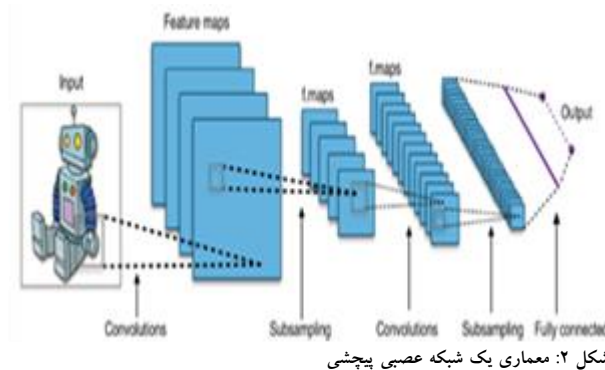
در طول دهه گذشته مطالعاتی با هدف آسیب‌شناسی خودکار دهان و دندان انجام شده‌است.^{۲۰-۲۴} پیش از ورود به مرحله تشخیص آسیب، برای به ثمر رساندن پیش‌پردازش تصاویر، لازم است تفکیک هر دندان از لثه یا قطعات مجاور به خوبی انجام شود تا آسیب در محل درست، تشخیص داده شود. همچنین از آنچه از این پیاده‌سازی‌ها به دست می‌آید می‌توان برای ثبت خودکار اطلاعات در پرونده‌های دیجیتال بیماران برای حفظ تاریخچه دندان و برنامه‌ریزی درمان استفاده کرد.^{۲۰}

با هدف تشخیص دندان، دو مطالعه^{۲۵،۲۶} به صورت اختصاصی روش‌های بخش‌بندی در سطح پیکسل را براساس روش‌های بینایی رایانه، مانند آستانه‌گذاری مبتنی بر هیستوگرام^{۱۸} پیشنهاد کرده‌اند. در

در استفاده از شبکه‌های عصبی پیچشی نسبت به بقیه رویکردهای دسته‌بندی تصاویر (Image classification) به میزان کمتری نیاز به پیش‌پردازش داده‌ها است. این امر به معنی آن است که شبکه معیارهایی را فرا می‌گیرد که در دیگر رویکردها به صورت دستی فرا گرفته می‌شوند. این استقلال از دانش پیشین و دستکاری‌های انسانی در شبکه‌های عصبی پیچشی، یک مزیت اساسی است.^{۱۵}

ساختار شبکه‌های پیچشی از فرایندهای زیستی قشر بینایی گربه الهام گرفته شده‌است. این ساختار به گونه‌ای است که تک نورون‌ها تنها در یک ناحیه محدود به تحریک پاسخ می‌دهند که به آن ناحیه پذیرش گفته می‌شود.^{۱۶}

لایه‌های پیچشی بررسی اولیه را روی ورودی اعمال می‌کنند، سپس نتیجه را به لایه بعدی می‌دهند (شکل ۲).^{۱۷} این پیچش در واقع پاسخ یک تک نورون را به یک تحریک دیداری شبیه‌سازی می‌کند. نواحی پذیرش نورون‌های مختلف به صورت جزئی با هم همپوشانی دارند، تا تمام ویژگی‌ها بررسی شوند و هیچ اطلاعاتی از توجه شبکه پنهان نماند.^{۱۶}



شکل ۲: معماری یک شبکه عصبی پیچشی

تشخیص دندان، برای تجزیه و تحلیل تصاویر و شروع شبیه‌سازی آنچه در ذهن درمانگر اتفاق می‌افتد، ابتدا سیستم به شناسایی کلیت تصویر (بدون آنکه تشخیص دهد چه چیزی را شناسایی می‌کند) می‌پردازد. این نوع شناسایی به وسیله جدا کردن قطعات موجود در تصویر اتفاق می‌افتد. در واقع سیستم پس از تشخیص مرز قطعات، آنها را به عنوان چند شی مجزا در نظر می‌گیرد. در تصویر OPG پس

بیشتری در بر داشته است، اما می‌توان این کاهش دقت را به علت جزئی‌نگری سیستم در بررسی داده موجود دانست. به‌عنوان مثال در مطالعه‌ای که مبتنی بر تصاویر رادیوگرافی پانورامیک دندانانی انجام شد، دندان، فیکسچر ایمپلنت و تاج با استفاده از الگوریتم R-CNN به ترتیب با نرخ‌های ۹۶/۷٪، ۴۵/۱٪ و ۶۰/۹٪ تشخیص داده شدند.^{۲۹} علاوه‌براین دقت، تشخیص ویژگی و دقت الگوریتم شماره‌گذاری دندان با استفاده از شبکه عصبی پیچشی و اکتشافی به ترتیب ۸۴/۲٪ و ۷۵/۵٪ بود.

مسئله‌ای که سیستم R-CNN به‌طور ویژه در آن کاربرد دارد، مکان‌یابی اشیا (Object detection) در تصویر است. تشخیص مرزها و تفکیک قطعات، برای متمایز کردن دندان از سایر قطعات کافی نیست، لذا استفاده از نوعی CNN که پس از بررسی تعداد زیاد دادگان، بتواند به دسته‌بندی اشیا از نظر شباهت پردازد، ضروری به‌نظر می‌رسد.^{۲۹}

همه موارد ذکر شده به جهت آموزش‌دهی صحیح به سیستم، احتیاج به داده‌ای دارد که به‌صورت برچسب‌گذاری شده در اختیار سیستم قرار گیرد و کامپیوتر با تطبیق برچسب‌ها و ویژگی‌های تصاویر، فرآیند یادگیری را آغاز کند.

مطالعاتی با هدف طراحی سیستم‌هایی برای برچسب‌گذاری خودکار تصاویر و بدون حضور متخصص انجام شده است. مطالعه‌ای که در سال ۲۰۲۲ انجام شد، بالاترین میزان دقت را در تفکیک دندان‌ها در سیستم‌هایی که تنها انسان نظارت‌گر برچسب‌گذاری سیستم است را داراست.^{۱۹}

با این‌حال رکورد دار شبکه‌های عصبی پیچشی در بحث دقت، مدل یادگیری عمیق YOLOv3 است که به‌دقت ۹۹/۳۳٪ دست یافت و از نظر دقت و جهانی بودن، بهتر از مدل‌های پیشرفته موجود عمل کرد.

YOLO مخفف عبارت (You only look once)، به معنای "شما فقط یک‌بار به تصویر نگاه می‌کنید" است. در واقع، این عبارت به همان قابلیت سیستم بینایی انسان اشاره دارد که با یک نگاه عمل تشخیص اشیا را انجام می‌دهد، بنابراین، سیستم تشخیص اشیا YOLO با هدف ارائه روشی مشابه کارکرد سیستم بینایی انسان طراحی شده است. ساختار کلی الگوریتم YOLO در حال تفکیک سه شی متفاوت در تصویر (شکل ۴).^{۳۰}

نهایت آنها دندان‌ها را با حساسیت به ترتیب ۰/۹۴ و ۰/۸۸ تشخیص دادند. البته با توجه به اینکه این مطالعات در سال ۲۰۰۹ و ۲۰۱۰ انجام شده‌اند این روش‌ها و سایر روش‌هایی که مبتنی بر پردازش تصاویر که در این حوزه انجام شده‌اند، همگی زیرمجموعه بینایی ماشین سنتی به حساب می‌آیند. لذا با استناد به نتیجه‌ی مطالعه‌ای که با هدف مشابه و در سال ۲۰۲۱ و با استفاده از دو نوع شبکه عصبی پیچشی انجام شد، می‌توان توانایی شبکه عصبی را در زمینه تشخیص دندان در تصاویر پانورامیک قدرتمندتر از روش‌های سنتی معرفی کرد.^{۲۷}

در این مطالعه از یک شبکه عصبی مصنوعی Matterport Mask RCNN در تشخیص اشیا استفاده شده است. همچنین یک شبکه ResNet101 به‌منظور طبقه‌بندی و شماره‌گذاری استفاده شده است. البته قابل ذکر است که در هر دو مورد از نوع شبکه‌های عصبی پیچشی که پیش‌تر به تحلیل آنها پرداخته شده است، استفاده شده است. معماری این مدل به دقت ۹۹/۲۴٪ در تشخیص دندان و ۹۳/۸۳٪ در شماره‌گذاری دندان با شرایط مختلف بهداشت دهان و دندان دست یافته است که به‌طور واضح خطای کمتری نسبت به مدل سنتی دارد.

نتایج مطالعه‌ای دیگر، نشان می‌دهد که با پیاده‌سازی الگوریتم‌های اکتشافی می‌توان نتایج حاصل از شبکه‌های عصبی عمیق پیچشی همانند RCNN را بهبود بخشید. در این مطالعه نتایج الگوی مبتنی بر یادگیری بیش از چهار درصد بهبود یافته است و دقت تا آستانه ۹۵/۸۷٪ بهبود یافته است.^{۲۸}

به‌صورت کلی‌تر می‌توان مدعی شد الگوریتم‌های اکتشافی می‌توانند دقت الگوریتم‌های مبتنی بر یادگیری عمیق را بهبود بخشند. این الگوریتم، روشی برای حل مسائلی است که راه‌حل‌های کلاسیک در آنها بسیار کند عمل می‌کنند یا به‌عنوان راه‌حل تقریبی برای مسائلی که راه‌های کلاسیک نمی‌توانند برای آنها جواب دقیقی پیدا کنند، پیشنهاد می‌شوند.

همچنین راه‌حل هیوریستیک مسئله از انعطاف‌پذیری بیشتری نسبت به راه‌حل دقیق آن برخوردار است که در نتیجه اجازه می‌دهد شروطی از مسئله که شبیه‌سازی آنها در حالت عادی سخت است، در راه‌حل هیوریستیک مدل شوند. البته پیش از آن مطالعاتی مشابه و با کارگیری الگوریتم‌های مکاشفه‌ای انجام شده است که نتایج خطای

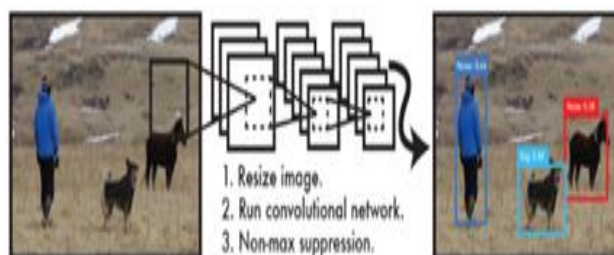
یک برچسب مشترک هستند دارای ویژگی یکسان یا مشابه می‌باشند.^{۱۳} با توجه به اینکه این معنابخشی در تصاویر به چه صورتی و با چه نسبتی صورت گیرد، فرآیندهای پیاده‌سازی شده، به دو نوع بخش‌بندی نمونه‌ای و بخش‌بندی معنایی متمایز می‌شوند.^{۱۹}

در بخش‌بندی نمونه‌ای المان یا شیء شناسایی شده در تصویر جداسازی می‌شود و شبکه هیچ تمایزی بین المان‌های شناسایی شده قائل نمی‌شود.^{۳۱} به‌طور مثال، تمام انسان‌های موجود در یک تصویر، به‌عنوان یک موجودیت و شیء دسته‌بندی می‌شوند و پس‌زمینه تصویر به عنوان یک موجودیت دیگر در نظر گرفته می‌شود.^{۲۱} اما در بخش‌بندی معنایی، موارد شناسایی شده در مرحله اول نیز معنابخشی می‌شود و سیستم کمی‌هوشمندانه‌تر عمل می‌کند و تفکری نزدیک‌تر به تفکر انسان دارد، چرا که اشیا دارای مفهوم متفاوت را در یک طبقه قرار داده و برای هر قطعه از تصویر، نمونه‌ای از موجودیتی که به آن تعلق دارد را می‌یابد.^{۳۳} به‌طور مثال، هر دندان در یک تصویر، به‌عنوان یک موجودیت مجزا بخش‌بندی می‌شود.

در این حوزه برای بخش‌بندی تصاویر OPG، ماشین با استفاده از شبکه‌های عصبی مختلف که اغلب بر پایه شبکه‌های عصبی پیچشی می‌باشند، تصاویر را به قسمت‌های کوچکتری از المان‌های شناسایی شده تقسیم‌بندی می‌کند و بنا به هدف مشخص، مطالعه آنها را آگاهانه‌تر از سر می‌گیرد.^{۱۰، ۲۲، ۲۳ و ۲۷}

در مطالعه‌ای که در سال ۲۰۲۱ انجام شد، همانطور که در شکل ۵ آمده است با استفاده از Mask R-CNN ابتدا دندان‌ها در جعبه‌های مرزی (Bounding Box) قرار داده شدند تا هر دندان از دندان مجاور تفکیک شود. سپس با استفاده از همین شبکه، دندان‌ها با بخش‌بندی معنایی تفکیک شدند. این مطالعه در نهایت با کاهش زمان تفکیک از ۶۰۰ ثانیه (تشخیص دانشجو یا متخصص) به سه ثانیه (تشخیص خودکار) توانست رکورد کاهش ۲۰۰ برابری زمان و تطبیق ۹۹/۰۲٪ را در این روش ثبت کند.^۹

در سال ۲۰۲۲ نیز مطالعه‌ای با هدف بخش‌بندی معنایی و تفکیک همه دندان‌های موجود در فک انجام شد. این مطالعه دو تفاوت عمده با مطالعه‌ی قبلی داشت. شبکه عصبی‌های استفاده شده در آن از انواع U-NET و جایگزینی روش «نقاط مرکزی» (Objects as Points) به جای «جعبه‌های مرزی» بود.^{۱۲ و ۳۳} نقاط مرکزی رویکردی برای ساده کردن روش جعبه‌های مرزی است.



شکل ۴: ساختار کلی الگوریتم YOLO در حال تفکیک سه شیء متفاوت در تصویر

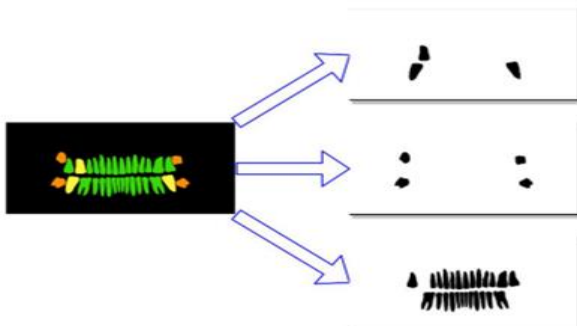
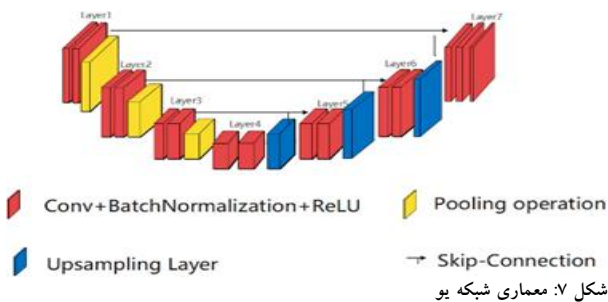
این نگرش سه ویژگی منحصر به فرد دارد، اول آنکه بسیار سریع است. دوم آنکه برای پیش‌بینی تشخیص، به صورت کلی (Global) به تصویر نگاه می‌کند. YOLO به شدت تعمیم‌پذیر است و در مقابل حوزه‌های جدید و یا داده‌های ورودی غیرمنتظره با احتمال کمتری نسبت به بقیه سیستم‌ها با شکست مواجه می‌شود. با توجه به آنکه ممکن است این الگوریتم در تشخیص اشیای ریزتر ضعیف عمل کند، در تشخیص و تفکیک دندان‌ها که نیازی به تفکیک موارد بسیار ریز نیست، دقت بالایی را به ثبت رسانده است.^{۳۲}

همان‌طور که پیش‌تر ذکر شد، مطالعات دیگری نیز با هدف مشابه در این حوزه انجام شده‌اند که از مدل‌های یادگیری ماشین استفاده نکرده‌اند. نتایج حاصل از این مطالعات ثابت می‌کند که تکنیک‌های سنتی بینایی کامپیوتر و پردازش تصویر به نسبت الگوریتم‌هایی که در آن سیستم به یادگیری آنچه در دسترس دارد می‌پردازد، از دقت کمتری در ارائه نتایج برخوردارند.

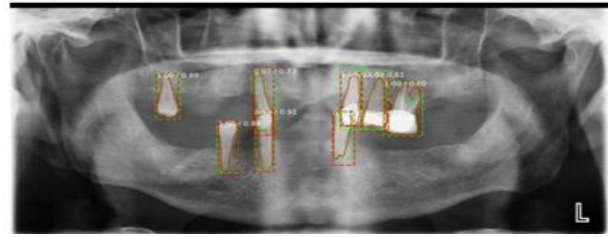
بخش‌بندی، پس از شناسایی کلیت تصاویر، مرحله‌ی بعدی که به شفافیت و فهم بهتر تصویر در نظر شبکه کمک می‌کند فرایند بخش‌بندی تصاویر است.^{۳۰ و ۳۱} در این مرحله خروجی مرحله‌ی پیش، برای سیستم هنوز مبهم بوده و صرفاً وجود اشیا درون تصویر به تایید سیستم رسیده و شناسایی شده است.

با انجام فرآیند بخش‌بندی، تلاش می‌شود، سیستم به‌صورت هوشمند به یادگیری هدفمند قسمت‌های مختلف تصویر و ویژگی‌های مربوط به اشیا تشخیص داده شده شامل مرزها، منحنی‌ها، موقعیت دقیق مکانی و ... پردازد. به عبارت دقیق‌تر، بخش‌بندی تصاویر، با اختصاص یک برچسب به هر پیکسل یا گروه پیکسل‌ها پیاده‌سازی می‌شود، به‌طوری که هر تعداد از پیکسل‌هایی که دارای

مشکل رایجی که می‌تواند در تجزیه و تحلیل خودکار رادیوگرافی پانورامیک دندان ظاهر شود، روی هم قرار گرفتن دندان‌ها است. این مشکل ممکن است به دلیل عدم انجام صحیح عکس‌برداری با اشعه ایکس و یا نزدیک بودن دندان‌های افراد به یکدیگر اتفاق بیفتد، پدیده‌ای که در اثر تغییر طبیعی در دندان فرد (تغییر دندان‌های شیری به دندان‌های دائمی) ایجاد می‌شود. این وضعیت می‌تواند باعث تقسیم‌بندی نادرست دندان شود که منجر به حالت نامطلوب دیدن دو یا چند دندان به عنوان یک نمونه می‌شود. مطالعه‌ای که در سال ۲۰۲۰ انجام شد به حل این مشکل پرداخت.^{۲۴} در اولین مرحله پس از پیش‌پردازش، شامل تقسیم تصویر بخش‌بندی شده اصلی به چندین تصویر، برابر با تعداد دسته‌ها بود، به طوری که تمام دندان‌های دارای دسته معنایی یکسان در یک تصویر قرار گرفتند.^{۲۵} سپس، هر تصویر با استفاده از یک مقدار آستانه، با رنگ دسته معنایی برای آن تصویر خاص، دودویی (باینری) شد، به این ترتیب پس‌زمینه سفید و پیش‌زمینه حاوی دندان‌ها، پروتزاها، ایمپلنت‌ها و غیره رنگ مشکی خواهد داشت (شکل ۸).

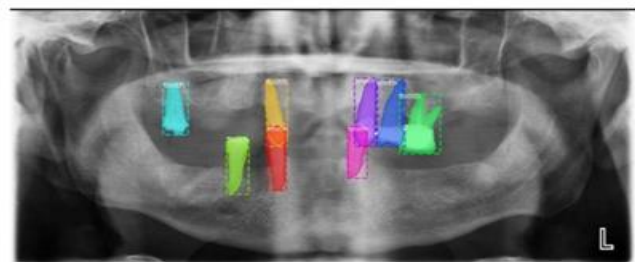


شکل ۸: تقسیم‌بندی دندان‌ها براساس کلاس معنایی و فرآیند دودویی کردن هر قطعه بندی تصویر



شکل ۵: تقسیم‌بندی نمونه‌ای دندان‌ها توسط جعبه‌های مرزی

این ایده هر شی را با یک نقطه در مرکز جعبه مرزی و سایر ویژگی‌ها مانند اندازه جسم، ابعاد، وسعت سه بعدی، جهت و حالت آن را به عنوان تابع رگرسیون ویژگی‌های تصویر نشان می‌دهد.^{۱۳} این موضوع بیانگر اینست که بخش‌بندی معنایی می‌تواند از طریق روش نقاط مرکزی و جعبه‌های مرزی (شکل ۶) با یک ساختار مناسب شبکه عصبی انجام پذیرد. قابل توجه است که شبکه عصبی انتخاب شده برای مطالعاتی با هدف بخش‌بندی در این حوزه اغلب از نوع U-NET است. معماری U-NET از معماری شبکه تماماً پیش‌بینی نشات می‌گیرد. ایده اصلی از ترکیب دو مسیر کاهنده و افزایشنده تشکیل شده است. داده‌های مسیر کاهنده با مسیر افزایشنده (که تعداد ویژگی‌های آن بسیار زیاد است) ترکیب می‌شود تا تصویر کیفیت بالایی با استفاده از اطلاعات مسیر کاهنده داشته باشد، سپس هر پیکسل به یک کلاس دسته‌بندی می‌شود و یک ویژگی به آن نسبت داده می‌شود.^{۳۴} در شکل ۷ این دو مسیر مشاهده می‌شود. در مسیر کاهشی (مسیر سمت چپ) تصویر به ویژگی‌های آن تجزیه و در مسیر افزایشی تصویر بازسازی می‌شود. در مسیر رو به بالا، فقط از ویژگی‌هایی از تصویر استفاده می‌شود که برای بخش‌بندی معنایی تصویر لازم است.^{۱۲}



شکل ۶: بخش‌بندی معنایی دندان‌ها در جعبه‌های مرزی

سرعت به R-CNN ها برتری دارد. در مرحله‌ی بعدی برای ایجاد شفافیت و تحلیل بهتر سیستم، صورت تصاویر بخش‌بندی شده‌اند و این کار به دو روش عمده انجام شده است. در برخی پیاده‌سازی‌ها، از دو شبکه تحلیلی استفاده شده است که نتیجه شبکه اول به‌عنوان ورودی شبکه دوم تعریف می‌شود.

ابتدا شبکه اول المان‌های موجود در تصویر را شناسایی می‌کند و سپس شبکه دوم، به بخش‌بندی و ایجاد تمایز بین المان‌های شناسایی شده می‌پردازد. این درحالی است که در بعضی از پیاده‌سازی‌ها تمامی این مراحل (شناسایی و بخش‌بندی) توسط یک شبکه، مانند U-Net، صورت گرفته و همانطور که انتظار تنها یک خروجی و با سرعت بیشتری به‌دست می‌آید.

لذا از این مطالعات برآورد می‌شود که در روش‌هایی که برای بخش‌بندی از شبکه‌های mask R-CNN و ... بهره گرفته شده است، علاوه بر زمان بیشتری که باید صرف شود، دقت پایین‌تری نیز در ایجاد تمایز بین خروجی‌ها به‌دست می‌آید.

در هر مرحله مقایسه‌های انجام شده نشانگر این است که در تشخیص، شبکه YOLO و به‌صورت خاص در بخش‌بندی معنایی، U-Net سرآمد شبکه‌های به کار گرفته شده در روش‌های مطالعات بررسی شده بوده است. با این حال طبق مطالعات و تحلیل‌های انجام‌شده، وابسته به ظرفیت‌ها و خروجی دلخواه و کارآمد، شبکه و روش مورد استفاده می‌تواند متغیر باشد که در این مرور به بعضی از آنها اشاره شد.^{۱۳و۳۰}

در ادامه یک عملیات ریخت‌شناسی روی تصاویر دودویی به دست آمده برای برخی کلاس‌های معنایی، به‌منظور جداسازی بهتر دندان‌ها پیاده‌سازی شد.

مقالات بررسی شده با هدف کلی بخش‌بندی تصاویر OPG در این مرور، به بخش‌بندی معنایی ناحیه‌ی دو فک^{۱۸و۲۹}، دندان‌ها^{۱۲و۹و۱۳و۳۰}، ضایعه‌ها^{۲۱و۲۳و۲۳}، ترمیم‌ها^{۲۱و۲۴} و نواحی با ویژگی‌هایی خاص^{۱۱} در ابعاد مختلف می‌پردازد. در حالت کلی همانند روش‌های تشخیص دندان، روش‌های به کار گرفته شده برای تقسیم‌بندی تصاویر نیز می‌تواند مبتنی بر یادگیری ماشین^{۱۱و۱۹و۳۴} یا بینایی کامپیوتر^{۱۸و۱۹} باشد. در بخش‌بندی اشیا موجود در تصویر، هر کدام از این دو روش، مسیرهای کلاسیک^{۱۰و۲۱} و به‌روز^{۲و۲۰و۲۱و۲۲و۲۳و۲۴و۲۵و۲۶و۲۷و۲۸} را پشتیبانی کرده و بخش‌بندی متفاوتی را در خروجی تولید می‌کنند.

جمع‌بندی، در این مرور مسیر تحلیل سیستم به‌صورت هوشمند در دو مرحله‌ی شناسایی (تشخیص) و بخش‌بندی (معنایی و نمونه) پیگیری شد. طبق مطالعات بررسی شده، ملاحظه می‌شود که برای مرحله‌ی تشخیص یا شناسایی اولیه المان‌های مدنظر در تصاویر رادیوگرافی سراسرنما بنا به دلایل ویژه‌ای، اغلب از شبکه عصبی پیچشی R-CNN و انواع آن بهره‌گیری می‌شود که در مقایسه با شبکه عصبی YOLO زمان بیشتر و دقت کمتری صرف می‌کند.

نتایج مطالعات بررسی شده نشان می‌دهد، شبکه عصبی YOLO که به‌صورت پیوسته خروجی‌ای متمایز ارائه می‌دهد، از لحاظ دقت و

References

1. J. G. Rossi, N. Rojas-Perilla, J. Krois, and F. Schwendicke, "Cost-effectiveness of artificial intelligence as a decision-support system applied to the detection and grading of melanoma, dental caries, and diabetic retinopathy," *JAMA Network Open*, vol. 5, no. 3, pp. e220269-e220269, 2022.
2. K. F. Hung, A. W. K. Yeung, M. M. Bornstein, and F. Schwendicke, "Personalized dental medicine, artificial intelligence, and their relevance for dentomaxillofacial imaging," *Dentomaxillofacial Radiology*, vol. 52, no. 1, p. 20220335, 2023.
3. S. Z. Seyyedsalehi and S. A. Seyyedsalehi, "A fast and efficient pre-training method based on layer-by-layer maximum discrimination for deep neural networks," *Neurocomputing*, vol. 168, pp. 669-680, 2015.
4. S. Z. Seyyedsalehi and S. A. Seyyedsalehi, "Why Dose Layer-by-Layer Pre-training Improve Deep Neural Networks Learning?," in *Handbook of Deep Learning Applications*: Springer, 2019, pp. 293-318.
5. H. Mohammad-Rahimi et al., "Deep learning for caries detection: a systematic review," *Journal of Dentistry*, vol. 122, p. 104115, 2022.
6. A. Müller, S. M. Mertens, G. Göstemeyer, J. Krois, and F. Schwendicke, "Barriers and enablers for artificial intelligence in dental diagnostics: a qualitative study," *Journal of Clinical Medicine*, vol. 10, no. 8, p. 1612, 2021.
7. F. Schwendicke, T. Golla, M. Dreher, and J. Krois, "Convolutional neural networks for dental image diagnostics: A scoping review," *Journal of dentistry*, vol. 91, p. 103226, 2019.
8. G. Silva, L. Oliveira, and M. Pithon, "Automatic segmenting teeth in X-ray images: Trends, a novel data set, benchmarking and future perspectives," *Expert Systems with Applications*, vol. 107, pp. 15-31, 2018.
9. A. Karaoglu, C. Ozcan, A. Pekince, Y. Yasa, B. Tekin, and D. Ozdemir, "Automatic dental segmentation module supported by artificial intelligence for dentistry students education," *Artificial*

- Intelligence Theory And Applications*, vol. 1, pp. 180-190, 2021.
10. K. Panetta, R. Rajendran, A. Ramesh, S. P. Rao, and S. Agaian, "Tufts dental database: a multimodal panoramic x-ray dataset for benchmarking diagnostic systems," *IEEE journal of biomedical and health informatics*, vol. 26, no. 4, pp. 1650-1659, 2021.
 11. A. H. Abdi, S. Kasaei, and M. Mehdizadeh, "Automatic segmentation of mandible in panoramic x-ray," *Journal of Medical Imaging*, vol. 2, no. 4, pp. 044003-044003, 2015.
 12. S. HELLÍ and A. Hamamci, "Tooth Instance Segmentation on Panoramic Dental Radiographs Using U-Nets and Morphological Processing," *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, vol. 10, no. 1, pp. 39-50, 2022.
 13. S. Lin, X. Hao, Y. Liu, D. Yan, J. Liu, and M. Zhong, "Lightweight deep learning methods for panoramic dental X-ray image segmentation," *Neural Computing and Applications*, vol. 35, no. 11, pp. 8295-8306, 2023.
 14. C. Huang, J. Wang, S. Wang, and Y. Zhang, "A review of deep learning in dentistry," *Neurocomputing*, p. 126629, 2023.
 15. S. Reza, S. A. Seyyedsalehi, and S. Z. Seyyedsalehi, "Modified deep attractor neural networks for variability compensation in recognition tasks," *Computers and Electrical Engineering*, vol. 99, p. 107776, 2022.
 16. G. W. Lindsay, "Convolutional neural networks as a model of the visual system: Past, present, and future," *Journal of cognitive neuroscience*, vol. 33, no. 10, pp. 2017-2031, 2021.
 17. S. Z. Seyyedsalehi and S. A. Seyyedsalehi, "Simultaneous learning of nonlinear manifolds based on the bottleneck neural network," *Neural processing letters*, vol. 40, pp. 191-209, 2014.
 18. D. V. Tuzoff *et al.*, "Tooth detection and numbering in panoramic radiographs using convolutional neural networks," *Dentomaxillofacial Radiology*, vol. 48, no. 4, p. 20180051, 2019.
 19. B. Silva *et al.*, "OdontoAI: A human-in-the-loop labeled data set and an online platform to boost research on dental panoramic radiographs," *arXiv preprint arXiv:2203.15856*, 2022.
 20. S. Vinayahalingam *et al.*, "Automated chart filing on panoramic radiographs using deep learning," *Journal of Dentistry*, vol. 115, p. 103864, 2021.
 21. T. Yeshua *et al.*, "Automatic detection and classification of dental restorations in panoramic radiographs," *Issues in Informing Science and Information Technology*, vol. 16, pp. 221-234, 2019.
 22. Y. E. Almalki *et al.*, "Deep learning models for classification of dental diseases using orthopantomography X-ray OPG images," *Sensors*, vol. 22, no. 19, p. 7370, 2022.
 23. J. Kim, H.-S. Lee, I.-S. Song, and K.-H. Jung, "DeNTNet: Deep Neural Transfer Network for the detection of periodontal bone loss using panoramic dental radiographs," *Scientific reports*, vol. 9, no. 1, p. 17615, 2019.
 24. M. P. Muresan, A. R. Barbura, and S. Nedevschi, "Teeth detection and dental problem classification in panoramic X-ray images using deep learning and image processing techniques," in *2020 IEEE 16th International Conference on Intelligent Computer Communication and Processing (ICCP)*, 2020: IEEE, pp. 457-463.
 25. P.-L. Lin, Y.-H. Lai, and P.-W. Huang, "An effective classification and numbering system for dental bitewing radiographs using teeth region and contour information," *Pattern Recognition*, vol. 43, no. 4, pp. 1380-1392, 2010.
 26. M. Hosntalab, R. Aghaeizadeh Zoroofi, A. Abbaspour Tehrani-Fard, and G. Shirani, "Classification and numbering of teeth in multi-slice CT images using wavelet-Fourier descriptor," *International journal of computer assisted radiology and surgery*, vol. 5, pp. 237-249, 2010.
 27. M. Prados-Privado, J. García Villalón, A. Blázquez Torres, C. H. Martínez-Martínez, and C. Ivorra, "A convolutional neural network for automatic tooth numbering in panoramic images," *BioMed Research International*, vol. 2021, 2021.
 28. A. Karaoglu, C. Ozcan, A. Pekince, and Y. Yasa, "Numbering teeth in panoramic images: A novel method based on deep learning and heuristic algorithm," *Engineering Science and Technology, an International Journal*, vol. 37, p. 101316, 2023.
 29. C. Kim, D. Kim, H. Jeong, S.-J. Yoon, and S. Youm, "Automatic tooth detection and numbering using a combination of a CNN and heuristic algorithm," *Applied Sciences*, vol. 10, no. 16, p. 5624, 2020.
 30. M. J. Shafiee, B. Chywl, F. Li, and A. Wong, "Fast YOLO: A fast you only look once system for real-time embedded object detection in video," *arXiv preprint arXiv:1709.05943*, 2017.
 31. D. Guo, Y. Pei, K. Zheng, H. Yu, Y. Lu, and S. Wang, "Degraded image semantic segmentation with dense-gram networks," *IEEE Transactions on Image Processing*, vol. 29, pp. 782-795, 2019.
 32. J. Yi, P. Wu, M. Jiang, Q. Huang, D. J. Hoepfner, and D. N. Metaxas, "Attentive neural cell instance segmentation," *Medical image analysis*, vol. 55, pp. 228-240, 2019.
 33. X. Zhou, D. Wang, and P. Krähenbühl, "Objects as points," *arXiv preprint arXiv:1904.07850*, 2019.
 34. J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431-3440.
 35. F. Schwendicke *et al.*, "Cost-effectiveness of artificial intelligence for proximal caries detection," *Journal of Dental Research*, vol. 100, no. 4, pp. 369-376, 2021.

An overview of the types of neural networks used in the examination of panoramic radiographic images: a narrative review

Abstract

Received: 9 Oct. 2024 Revised: 16 Oct. 2024 Accepted: 13 Nov. 2024 Available online: 21 Nov. 2024

Mahdieh Soltani M.Sc.^{1,2}
 Seyede Zohreh
 Seyedsalehi Ph.D.^{1*}
 Reyhane Mahdavi B.Sc.¹

1- Department of Biomedical Engineering, Faculty of Paramedical Sciences, Tehran Azad University of Medical Sciences, Tehran, Iran.

2- Department of Biomedical Engineering, Faculty of Engineering, Shahed University, Tehran, Iran.

With the rapid expansion of artificial intelligence across clinical disciplines, a variety of artificial neural networks (ANNs) have become indispensable tools for endowing computer systems with advanced analytical power. Dentistry, as an information-rich branch of medicine, routinely generates and must interpret large, complex datasets from imaging and diagnostic records. Consequently, researchers have increasingly directed their attention toward intelligent, automated techniques for analyzing dental data. This study therefore surveys and synthesizes the methods that have been applied to the intelligent and automated analysis of such data, highlighting the prevailing trends in current literature. The majority of the examined investigations relied on panoramic radiographic images of the teeth orthopantomograms (OPG) as their primary source material. Three overarching technical objectives repeatedly emerged: first, tooth diagnosis, meaning the reliable separation and identification of each individual tooth from its neighbors; second, sample segmentation, that is, the piece-by-piece analysis of visual information within the image; and third, semantic segmentation, namely, the contextual interpretation of information extracted from the radiograph. Depending upon which of these objectives was pursued, researchers selected different neural-network architectures and configurations. Across the reviewed corpus, input images were typically subjected to preprocessing steps such as normalization, noise reduction, and contrast enhancement before being supplied to a neural network for training, thereby preparing the data for subsequent machine interpretation. In several instances, the raw output produced by the neural network underwent additional post-processing, a stage designed to refine the preliminary results and enhance overall accuracy. The comparative analysis presented here concentrates on how effectively the various neural-network models fulfilled the three technical objectives described above. The surveyed articles reveal two dominant analytical approaches. In the intelligent problem-solving paradigm, convolutional neural networks (CNNs) overwhelmingly predominate. Conversely, in the automated paradigm, investigators favor classical, non-learning algorithmic techniques. Work employing ANNs consistently emphasizes image comprehension, segmentation, feature extraction, feature classification, network modeling, and careful variable tuning to promote effective learning that aligns with each study's stated objectives.

Keywords: artificial neural networks, oral and dental diagnostics, panoramic radiography.

* Corresponding author:
 Department of Biomedical Engineering, Faculty of Paramedical Sciences, Tehran Azad University of Medical Sciences, Tehran, Iran.
 Tel: +98-21-22006660
 E-mail:
 z.seyyedsalehi@iautmu.ac.ir

Copyright © 2024 Soltani et al. Published by Tehran University of Medical Sciences.

This work is licensed under a Creative Commons Attribution-Non-Commercial 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.

Tehran Univ Med J (TUMJ) 2024 December;82(9):715-23

<http://tumj.tums.ac.ir>