مقاله پژوهشی، صفحات ۱۶۳-۱۷۹
DOI: 10.22034/cmr.2026.8642.3327

تشخیص خودکار سرطان ریه از تصاویر هیستوپاتولوژی با روش یادگیری عمیق

سیده سمانه شجاعی لنگری[1*]، ابوالفضل بزازپور[2]

[1]گروه مهندسی پزشکی، پژوهشکده برق و فناوری اطلاعات، سازمان پژوهش های علمی و صنعتی ایران، تهران، ایران
[2]گروه GIS، دانشکده سنجش از دور و GIS، دانشگاه شهید بهشتی، تهران، ایران

تاریخ دریافت: ۱۴۰۴/۰۳/۲۰ تاریخ پذیرش: ۱۴۰۵/۰۵/۲۶

چکیده

سرطان یکی از عمده‌ترین علل مرگ و میر در جهان است و تشخیص زودهنگام آن نقش کلیدی در مدیریت مؤثر دارد. روش استاندارد تشخیص، بیوپسی و بررسی دستی اسلایدهای هیستوپاتولوژی توسط پاتولوژیست‌هاست که زمان‌بر، نیازمند تجربه بالا و مستعد خطا است. . ازاین‌رو، استفاده از سیستم‌های تشخیص رایانه‌ای مبتنی بر هوش مصنوعی می‌تواند به بهبود کارایی تفسیر تصاویر پزشکی کمک کند. روش‌های هوش مصنوعی توانایی در استخراج اطلاعات کمّی از تصاویر دیجیتال هیستوپاتولوژی را چند برابر کرده است و انتظار می‌رود که بتواند حجم کار برای کارشناسان انسانی را کاهش داده، صحت و پیوستگی گزارش‌های هیستوپاتولوژی را بهبود بخشد. در این تحقیق از مدل‌های مختلف یادگیری عمیق از جمله ResNet50، ResNet152v2 و Inception-v3 برای تشخیص سرطان ریه و طبقه بندی زیررده‌های آن از روی تصاویر هیستوپاتولوژی استفاده شده است. نتایج این تحقیق بر روی بانک داده LC25000 با استفاده از مفهوم یادگیری انتقالی، بیانگر پتانسیل بالقوه روش پیشنهادی در تشخیص این بیماری می‌باشد. در این میان، شبکه ResNet152v2 بهترین عملکرد را داشته که به صحت ۹۹.۶٪، دقت ۹۹.۶٪ و حساسیت ۹۹.۶٪ دست یافته است. اما پس از اعمال کنترل نشت داده و جداسازی دقیق تصاویر بیماران در مجموعه‌های آموزش و آزمون، این مقادیر به ترتیب به ۹۶.۹٪، ۹۶.۷٪ و ۹۶.۷٪ کاهش یافت که ارزیابی واقع‌بینانه‌تری از مدل ارائه می‌دهد. به‌طورکلی، یافته‌ها حاکی از ظرفیت بالای روش‌های یادگیری عمیق در طبقه‌بندی زیررده‌های سرطان ریه است. بااین‌حال، تعمیم‌پذیری نتایج به محیط بالینی نیازمند اعتبارسنجی بر روی مجموعه‌داده‌های مستقل و متنوع، شامل اسلایدهای کامل، در مطالعات آتی است.

واژگان کلیدی: تصاویر هیستوپاتولوژی، سرطان ریه، سیستم‌های تشخیص خودکار، شبکه عصبی کانولوشن، یادگیری عمیق

* نویسنده مسئول، پست الکترونیکی: s.shojaie@irost.ir

مقدمه

مقدمه را اینجا وارد کنید. سرطان به گروه پیچیده‌ای از بیماری‌های غیر واگیری اطلاق می‌شود که با تکثیر و گسترش کنترل نشده سلول‌های غیرطبیعی مشخص می‌شود. این سلول‌ها دستخوش تغییرات ژنتیکی و اپی‌ژنتیکی می‌شوند که به جهش‌ها معروف هستند، این جهش‌ها آنها را قادر می‌سازد از مکانیسم‌های کنترل رشد طبیعی فرار کنند و بدون کنترل به تکثیر خود ادامه

مقاله پژوهشی، صفحات ۱۶۳-۱۷۹
DOI: 10.22034/cmr.2026.8642.3327

دهند توسعه سرطان شامل یک تعامل پیچیده بین عوامل ژنتیکی، اپی‌ژنتیکی و محیطی است [۱۳] . در نتیجه سلول‌های سرطانی تجمع یافته و تومورها را تشکیل می‌دهند یا به بافت‌ها و اندام‌های مجاور حمله می‌کنند. اگر این تکثیر کنترل نشود، ممکن است منجر به مرگ بیمار شود. براساس گزارش‌های سازمان بهداشت جهانی، سرطان دومین عامل مرگ‌ومیر در سطح جهان است و در سال ۲۰۱۸، حدود ۹/۶ میلیون مرگ بر اثر سرطان گزارش شده است. بنابراین سرطان یک معضل جدی است که سلامت تمامی جوامع بشری را تحت تأثیر قرار می‌دهد. متأسفانه سرطان یک بیماری متنوع در سطح بافت است و این تنوع یک چالش بزرگ برای تشخیص دقیق آن و به دنبال آن درمان اثربخش می‌باشد.

سرطان ریه یکی از شایع‌ترین علل مرگ و میر برای بیماران سرطانی در سراسر جهان است و هر ساله میلیون‌ها نفر را به مرگ می‌کشاند. در قرن بیستم، به دلیل افزایش مصرف سیگار و تنباکو توسط جمعیت عمومی، مرگ و میر ناشی از سرطان ریه افزایش یافت. در حال حاضر، در برخی کشورها، به دلیل محدودیت‌های اعمال شده برای مصرف تنباکو، این آمار کاهش یافته است. اما در اکثر جوامع، این مشکل به سرعت در حال افزایش است. آمارهای ثبت شده نشان می‌دهند که میزان بروز بیماری تقریباً با میزان مرگ بیماران ناشی از آن تقریباً برابر است. در سنین پایین‌تر از ۵۰ سال، میزان بروز سرطان ریه نسبتاً پایین است و در مردان و زنان برابر است، اما پس از این سن، به ویژه پس از ۶۰ سالگی، میزان بروز آن به سرعت افزایش می‌یابد و با افزایش سن، این میزان در مردان بیشتر از زنان می‌شود. به طور کلی، سرطان ریه عامل اصلی مرگ و میر ناشی از سرطان در مردان و دومین علت مرگ در زنان است [۲۴].

نتایج درمان سرطان ریه در مقایسه با سرطان سینه و پروستات نامطلوب‌تر است، زیرا سرطان سینه و پروستات معمولاً در مراحل اولیه بیماری تشخیص داده می‌شوند و اقدامات مناسب برای درمان آنها در نظر گرفته می‌شود. در حالی که اکثر بیماران مبتلا به سرطان ریه در مرحله‌ای از بیماری تشخیص داده می‌شوند که به متاستاز رسیده است. بیشترین ضایعات سرطان ریه نیز در مراحل پیشرفته تشخیص داده می‌شود، به همین دلیل، پزشکان زمان کافی برای درمان بیماران را ندارند. برای سرطان ریه، معمولاً میزان بقای ۵ ساله حدود ۱۵٪ است. با این حال، زمانی که بیماری در مراحل اولیه تشخیص داده می‌شود، میزان بقا به طور چشمگیری افزایش می‌یابد [۱۰].

در حال حاضر، تشخیص اولیه سرطان ریه مبتنی بر تجزیه و تحلیل تصاویر CT توسط پاتولوژیست می باشد و تایید تشخیص نیاز به نمونه گیری های تهاجمی دارد و سپس تصاویر نمونه (هیستوپاتولوژی) توسط پاتولوژیست مورد تجزیه و تحلیل قرار می‌گیرد. در سی تی اسکن ریه از اشعه ایکس برای بررسی ناهنجاری های موجود استفاده می شود. این روش سریع ، بدون درد ، غیر تهاجمی و دقیق است و از آنجایی که قادر به تشخیص گره های بسیار کوچک در ریه می باشد برای تشخیص سرطان ریه در اولین مراحل تجویز می‌شود. در صورتیکه نتایج تصویر برداری سی‌تی دوپهلو باشد، از روش‌های تهاجمی مثل برونکوسکوپی استفاده می‌شود. بیوپسی از نواحی تومور از طریق برونکوسکوپی، روش استاندارد طلایی محسوب می‌شود ولی در مواردی که تومور در نواحی محیطی ریه قرار دارد و یا بیمار با خطر خونریزی مواجه است، انجام آن مقدور نمی‌باشد و در این صورت تشخیص قطعی مبتنی بر مشاهده تصاویر تهیه شده از نمونه بیوپسی شده به صورت جراحی یا لاپاراسکوپی توسط پاتولوژیست می باشد. در این میان

مقاله پژوهشی، صفحات ۱۶۳-۱۷۹
DOI: 10.22034/cmr.2026.8642.3327

تجربه پاتولوژیست در تشخیص صحیح بسیار مهم است. در بسیاری از موارد به دلیل تاخیر یا عدم ارسال صحیح نمونه از بیمارستان محل بیوپسی به آزمایشگاه و یا عدم تجربه پاتولوژیست منجر به تشخیص غلط می‌شود. به عبارتی تشخیص منفی و مثبت کاذب هر دو باعث مرگ بیمار می شود. منفی کاذب باعث عدم درمان و متاستاز سرطان و نهایتا مرگ بیمار و مثبت کاذب باعث اعمال هزینه‌ی گزاف و همچنین دریافت داروی شیمی درمانی که خود می تواند عامل ایجاد سرطان باشد می‌گردد. متاسفانه طبق گزارشات علمی، حدود ۱۵-۳۵ درصد احتمال تشخیص غلط سرطان ریه در جوامع مختلف وجود دارد. لذا گسترش سیستم‌های خودکار کامپیوتری جهت کاهش خطای انسانی و افزایش دقت تشخیص بسیار مورد نیاز می باشد.

در این تحقیق، یک روش تشخیصی جهت غربالگری تصاویر هیستوپاتولوژی ندول‌های ریوی با استفاده از یک طبقه‌بندی‌کننده مبتنی بر شبکه‌های عصبی عمیق پیشنهاد می‌شود. عملکرد روش پیشنهادی بر اساس شاخص‌های کمی ارزیابی شامل صحت (Accuracy)، دقت (Precision)، حساسیت (Sensitivity) و سایر معیارهای مبتنی بر ماتریس درهم‌ریختگی گزارش شده است. نتایج به‌دست‌آمده نشان می‌دهد که مدل پیشنهادی در مقایسه با روش‌های مورد بررسی، عملکرد رقابتی و مطلوبی در شاخص‌های ارزیابی ارائه می‌کند.

از زمان معرفی تجاری تصویربرداری کامل اسلایدها (WSI) ، آسیب شناسی دیجیتالی به یک میدان پر رونق در دنیا تبدیل شده است. توانایی دیجیتالی کردن کل یک اسلاید شیشه ای، باعث ایجاد کاربردهای بالینی، آموزشی و تحقیقاتی شده است. آسیب شناسی دیجیتال، پزشکان و آسیب شناسان را قادر می سازد تا با شفافیت و سازگاری ، سرعت و ارزیابی از راه دور و همکاری سریع و درنتیجه بهبود کارایی و بهره وری به فعالیت‌های تشخیصی خود ادامه دهند [۱۳]،[۱۰]. آینده آسیب شناسی دیجیتال می‌تواند در نهایت شامل تحقیقات پیشرفته تشخیصی با کمک سیستم های تشخیصی هوش مصنوعی و پیشرفت چشمگیر پزشکی تشخیصی باشد.

پاتولوژی محاسباتی به کاربردهای هوش مصنوعی در تحلیل تصویر مبتنی بر کامپیوتر در هیستوپاتولوژی اشاره دارد که در گذشته ایجاد و به طور گسترده پذیرفته شده است. روش‌های پاتولوژی محاسباتی قادرند اطلاعات را از فنوتیپ هیستولوژیکی استخراج کنند و در نتیجه مجموعه‌ای از کاربردهای تحقیقاتی و تشخیصی جدید را ممکن می‌سازند [۲۳]. این توانایی، نتیجه پیشرفت‌های فناوری سریع در حوزه‌های هیستوپاتولوژی و تشخیص مولکولی، به همراه نوآوری‌های روش‌شناسی در بینایی ماشین و بیوانفورماتیک در دهه‌های گذشته است. بررسی ویژگی‌های متنوع تومور در سطح مولکولی به کمک فناوری‌های جدید با قابلیت تولید توصیف دقیق‌تر و بیشتر از تومورها و محیط میکرو آنها، ممکن شده است. بسیاری از اطلاعات زیستی و بالینی با اهمیت مانند بیومارکرهای ایمنوتراپی سرطان و قابلیت پیش‌بینی بیماری، می‌توانند با استفاده از ابزارهای بیوانفورماتیک و علوم داده استخراج شوند. در مقابل، بررسی تجسمی فنوتیپ‌های سرطانی هیستوپاتولوژیکی، که برای تشخیص تومورهای جامد ضروری است، به طور سنتی توسط کارشناسان انسانی و بدون کمک فناوری تنها از طریق یک میکروسکوپ نوری، انجام می‌شود. بررسی تجسمی بخش‌های بافتی در اسلایدهای هیستوپاتولوژیکی برای تشخیص تومورهای جامد ضروری است. یک بخش هیستوپاتولوژیکی معمولا شامل صدها هزار سلول است. بنابراین، علاوه بر کمک به کار تشخیص روزمره از طریق اتوماسیون، اندازه‌گیری کمی

DOI: 10.22034/cmr.2026.8642.3327
مقاله پژوهشی، صفحات ۱۶۳-۱۷۹

مبتنی بر کامپیوتر از این نمونه‌ها می‌تواند مقدار زیادی اطلاعات علمی و پزشکی ارزشمند را تولید کند.

مطالعات اخیر مروری نشان داده‌اند که شبکه‌های عصبی پیچشی (CNNs) رایج‌ترین معماری‌های به کار گرفته شده در تحلیل تصاویر هیستوپاتولوژی سرطان ریه هستند [۱۹،۲۰،۲۱،۳]. این مدل‌ها توانایی قابل توجهی در پیش‌بینی تغییرات مولکولی محرک سرطان (مانند EGFR، ALK و KRAS) از روی تصاویر رنگ‌آمیزی شده با H&E دارند. به عنوان مثال، متاآنالیزی که بر روی ۲۳ مطالعه معتبر انجام شده، حساسیت و ویژگی ٪۸۴ و ٪۸۵ را برای پیش‌بینی تغییرات ALK گزارش کرده است [۲۰]. همچنین، کاربرد این مدل‌ها در طبقه‌بندی زیرگروه‌های بافتی سرطان ریه (آدنوکارسینوما در مقابل کارسینوم سلول سنگفرشی) با دقت‌های بالایی همراه بوده است [۱۹].

شجاعی و همکاران طرح در تحقیقی با همکاری دانشگاه تربیت مدرس و انستیتو بولتزمان اتریش، با بکار گیری الگوریتم‌های یادگیری عمیق بر روی تصاویر سلول‌های سرطانی ریه کشت داده شده در بستر میکروفلویدیک (به منظور شباهت هر چه بیشتر به بافت‌های طبیعی) به نتایج قابل توجهی در تشخیص بین سلول های نرمال و سرطانی (با دقت ٪۹۹.۷۷) و حتی طبقه بندی بین زیرشاخه‌های سرطان ریه (با صحت ٪۹۸.۳۷) دست یافتند [۱۲].

کاناواتی و همکاران از یک شبکه کانولوشنی (مدل EfficientNet-B3) برای طبقه بندی سلول های نرمال و سرطانی از روی ۳۷۰۴ تصویر هیستوپاتولوژی ریه (جمع آوری شده از مدیکال سنتر کیوشو، توکیو) استفاده کرده و به نتایج قابل قبولی دست یافتند [۱۵].

تراموتو و همکاران یک طرح طبقه بندی خودکار برای تشخیص انواع سرطان ریه (شامل adenocarcinoma, squamous cell carcinoma, and small cell carcinoma) از روی تصاویر میکروسکوپیک به روش یادگیری عمیق کانولوشنی ارائه کردند و به صحت ٪۷۱ دست یافتند که قابل قیاس با نتایج گزارش شده توسط پاتولوژیست‌ها روی داده مربوطه بود [۲۷].

کودری و همکاران با بکارگیری شبکه یادگیری عمیق مدل inceptionv3 روی تصاویر هیستوپاتولوژی از بانک داده در دسترس TCGA به نتایج خوبی در طبقه بندی انواع adenocarcinoma و squamous cell carcinoma از سرطان ریه و همچنین بافت نرمال دست یافتند که قابل قیاس با نتایج گزارش شده توسط پاتولوژیست ها روی این مجموعه داده می باشد [۷] .

وی و همکاران به کمک شبکه عمیق ResNet روشی جهت تشخیص خودکار نوع adenocarcinoma از سرطان ریه از روی تصاویر هیستولوژی بافت جراحی شده (۱۴۳ اسلاید) ارائه کردند که به نتیجه کاپا اسکور ۰.۵۲ با توافق ٪۶۶.۶ با سه پاتولوژیست دست یافتند [۲۸].

چن و همکاران با بکارگیری یک روش نوین آموزش در طراحی شبکه‌های عمیق یادگیری موفق به تشخیص دو نوع سرطان ریه (adenocarcinoma و squamous cell carcinoma) از روی تصاویر هیستوپاتولوژی به صورت whole-slide شدند [۶]. طبق گزارشات چاپ شده از این تحقیق، روش پیشنهادی آن‌ها روی ۹۶۶۲ تصویر هیستوپاتولوژی ریه با نتیجه سطح زیر منحنی ۰.۹۵ و ۰.۹۴ برای دو نوع نامبرده سرطان ، عملکرد بهتری نسبت به روش‌های یادگیری متداول دارد.

لو و همکاران نیز روشی بر مبنای شبکه عصبی با یادگیری عمیق به روش weak-supervised برای طبقه‌بندی زیر شاخه‌های دو نوع سرطان ریه (non-small-cell و renal cell carcinoma) و همچنین

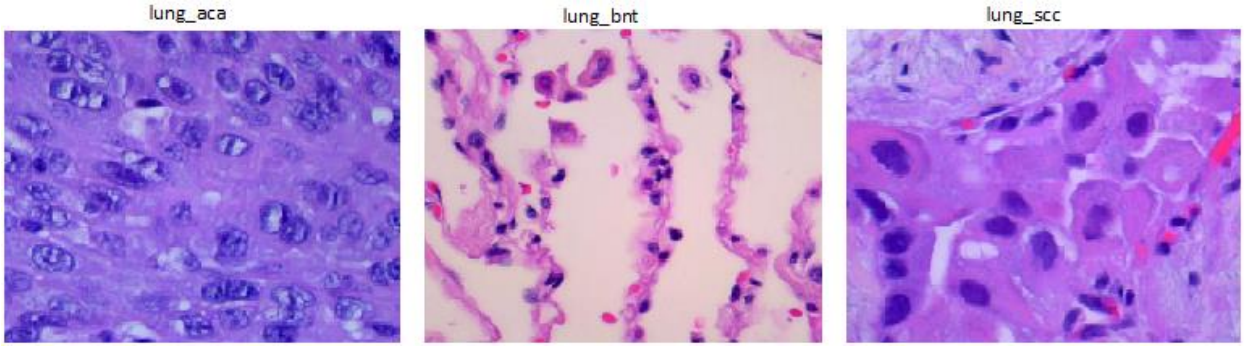
مقاله پژوهشی، صفحات ۱۶۳-۱۷۹
DOI: 10.22034/cmr.2026.8642.3327

تشخیص متاستاز گره‌های لنفی ارائه کرده و به نتایج قابل توجهی روی بانک داده TCGA دست یافتند [۱۷]. آن‌ها همچنین عملکرد روش پیشنهادی خود را روی بانک داده‌ی دیگری تست کرده و نشان دادند که الگوریتم آن‌ها قابل تعمیم برای کاربردهای کلینیکی می‌باشد.

بانک داده

بانک داده LC25000

در این پژوهش از مجموعه داده تصاویر هیستوپاتولوژیکی سرطان ریه و روده بزرگ به نام LC25000 استفاده کردیم تا مدل پیشنهادی خود را ارزیابی کنیم [۴]. این مجموعه داده توسط آندرو بورکوفسکی و همکارانش در بیمارستان جیمز هاستپیتال در تامپا، فلوریدا، تهیه شده است و شامل ۲۵,۰۰۰ تصویر به دو نوع سرطان روده بزرگ و سه نوع سرطان ریه تقسیم می‌شود. تصاویر به طور مساوی بین پنج کلاس توزیع شده‌اند، به این معنی که مجموعه داده متعادل است و هر نوع شامل ۵۰۰۰ تصویر می‌باشد. در این تحقیق، تنها از داده های سرطان ریه استفاده شد که شامل سه کلاس بافت خوش خیم ریه (lung_bnt)، آدنوکارسینوم ریه (lung_aca)، و کارسینوم سلول سنگفرشی (lung_scc) می‌باشد. تصویر ۱ نمونه‌هایی از مجموعه داده LC25000 برای سه کلاس نامبرده را نشان می‌دهد.



تصویر ۱: نمونه تصاویر هیستوپاتولوژی از بانک داده LC25000

بانک داده LC25000-clean

بر اساس مقاله اصلی معرفی بانک داده LC25000، این بانک داده از ۷۵۰ تصویر پایه (۲۵۰ تصویر برای هر کلاس ریه و ۲۵۰ تصویر برای هر کلاس روده) از طریق اعمال تبدیلات چرخش (تا ۲۵ درجه) و برگرداندن افقی و عمودی تولید شده است. هر تصویر پایه به طور میانگین به ۲۰ تصویر افزایش‌یافته تبدیل شده است. به عبارت دیگر، در نسخه اصلی بانک داده، این ۲۰ تصویر که از یک تصویر پایه تولید شده‌اند، به طور تصادفی در کل دیتاست پخش شده‌اند، بدون اینکه برچسب‌گذاری شوند که به یک خانواده تعلق دارند. این مسئله می‌تواند منجر به نشت داده (Data Leakage) شده و دقت‌های گزارش‌شده را به طور مصنوعی افزایش دهد.

به منظور جلوگیری از نشت داده ناشی از افزایش تصنعی تصاویر پایه، از نسخه پالایش‌شده دیتاست با نام LC25000-clean استفاده شد [۳]. این نسخه توسط Batchkala و همکاران در سال ۲۰۲۴ ارائه شده است. پژوهشگران یک خط لوله نیمه‌خودکار برای پاک‌سازی مجموعه داده ایجاد کردند که با استفاده از ویژگی‌های استخراج‌شده توسط مدل‌های بنیادی (foundation models) و سپس خوشه‌بندی (clustering) همه‌ی تصاویر تولیدشده از یک تصویر اصلی را جدا کرده و به صورت گروهی تفکیک می‌کند. این کار به این معنی است که همه‌ی نسخه‌های افزایشی یک تصویر در همان گروه قرار می‌گیرند و سپس تقسیم‌بندی آموزش/آزمون به گونه‌ای انجام می‌شود که گروه‌های مرتبط از هم جدا

مقاله پژوهشی، صفحات ۱۶۳-۱۷۹
DOI: 10.22034/cmr.2026.8642.3327

باشند، بنابراین نشت داده کاهش می‌یابد و ارزیابی عملکرد مدل واقعی‌تر می‌شود.

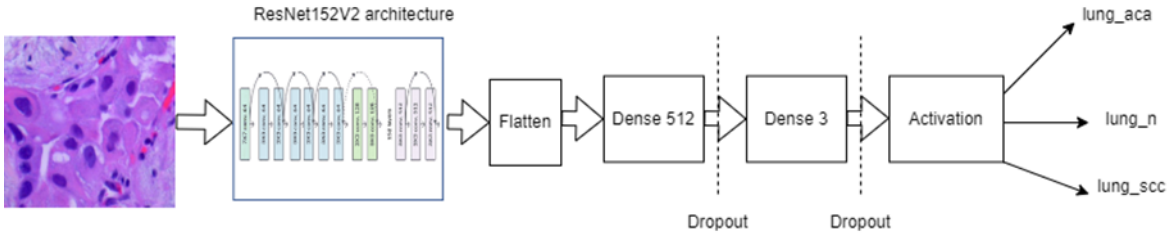
مواد و روش‌ها

این بخش ابتدا به معرفی شبکه‌های عصبی عمیق انتخاب شده در این پژوهش پرداخته و سپس پروسه‌ی یادگیری انتقالی اعمال شده بر داده‌ی مورد نظر را توضیح می‌دهد.

معماری ResNet

معماری ResNet یکی از شبکه‌های عصبی کانولوشنی پرکاربرد است که با معرفی اتصال‌های کوتاه (Skip Connections) مشکل فروپاشی گرادیان در شبکه‌های عمیق را مرتفع ساخته است [۱۴]. این ویژگی امکان انتقال مستقیم سیگنال ورودی به لایه‌های بالاتر را فراهم کرده و یادگیری مؤثر در شبکه‌های بسیار عمیق را میسر می‌سازد.

نسخه‌های مختلفی از این معماری مانند ResNet18، ResNet50 و ResNet101 برای کاربردهای گوناگون وجود دارند، اما در این پژوهش، پس از آزمایش‌های اولیه، معماری ResNet152v2 به عنوان مدل برتر انتخاب شد. این معماری شامل ۱۵۲ لایه عمیق با بلوک‌های باقیمانده (Residual Blocks) است و از تکنیک‌های بهینه‌سازی نظیر نرمال‌سازی دسته‌ای (Batch Normalization) بهره می‌برد که ضمن افزایش سرعت یادگیری، پایداری مدل را بهبود می‌بخشد. تصویر ۲ نمای کلی این معماری را نشان می‌دهد.



تصویر ۲: معماری شبکه پیشنهادی ResNet152v2

معماری Inception-v3

Inception-v3 یک معماری شبکه عصبی کانولوشنی پیشرفته از خانواده Inception است که در سال ۲۰۱۵ توسط محققان گوگل معرفی شد [۲۶]. ویژگی اصلی این معماری، استفاده از مدول‌های Inception است که با اجرای هم‌زمان چندین فیلتر کانولوشنی با اندازه‌های مختلف (۱×۱، ۳×۳، ۵×۵) به همراه max pooling، امکان استخراج ویژگی‌ها در مقیاس‌های گوناگون را فراهم می‌کند.

این شبکه با ۴۸ لایه عمیق‌تر از نسخه‌های پیشین خود بوده و با بهره‌گیری از تکنیک‌های کاهش ابعاد، کارایی محاسباتی را بهبود می‌بخشد. معماری Inception-v3 علاوه بر کاربرد گسترده در طبقه‌بندی تصاویر و تحلیل تصاویر پزشکی، در حوزه‌هایی مانند پردازش زبان طبیعی نیز مورد استفاده قرار می‌گیرد.

یادگیری انتقالی

یادگیری انتقالی به‌کارگیری دانش یک مدل از پیش آموزش‌دیده برای حل مسئله‌ای جدید اما مرتبط است. در این رویکرد، به‌جای آموزش شبکه از ابتدا، وزن‌های آموخته‌شده در یک مجموعه‌داده بزرگ به مسئله هدف منتقل می‌شوند تا در شرایط کمبود داده، تعمیم‌پذیری و پایداری مدل بهبود یابد.

در شبکه‌های عصبی کانولوشنی، لایه‌های ابتدایی معمولاً ویژگی‌های عمومی تصویر (مانند الگوهای لبه و بافت) را استخراج می‌کنند؛ ازاین‌رو می‌توان این لایه‌ها را

مقاله پژوهشی، صفحات ۱۶۳-۱۷۹
DOI: 10.22034/cmr.2026.8642.3327

به‌عنوان استخراج‌کننده ویژگی ثابت نگه داشت و تنها لایه‌های انتهایی را برای مسئله جدید بازآموزی (fine-tune) کرد. این راهبرد ضمن کاهش تعداد پارامترهای قابل آموزش، زمان همگرایی و نیاز به داده را نیز کاهش می‌دهد.

در این پژوهش، عملیات انتقال یادگیری با استفاده از معماری‌های از پیش آموزش‌دیده ResNet50، ResNet152V2 و Inception-v3 انجام شد. در هر مدل، لایه‌های اولیه به‌صورت ثابت حفظ شده و ۱۰ لایه انتهایی متناسب با مسئله حاضر بازآموزی شدند. تمرکز این مطالعه بر مقایسه عملکرد این معماری‌ها در چارچوب داده‌های مسئله مورد بررسی و تحلیل تأثیر راهبرد بازآموزی محدود لایه‌ها بر دقت نهایی مدل است.

جزئیات پیاده‌سازی و تنظیمات آموزشی

برای بهینه‌سازی مدل‌های آموزشی، از الگوریتم AdamW با نرخ یادگیری ۴-۱۰×۱ و ضریب کاهش وزن (Weight Decay) برابر با ۳-۱۰×۱ استفاده شد. تابع هزینه،CrossEntropyLoss همراه با هموارسازی برچسب‌ها (Label Smoothing) به میزان ۰.۱ انتخاب گردید که به کاهش بیش‌اطمینانی مدل کمک می‌کند. همچنین، به‌منظور جلوگیری از انفجار گرادیان، تکنیک Gradient Clipping به کار گرفته شد.

برنامه زمان‌بندی نرخ یادگیری به‌صورت دو مرحله‌ای تنظیم شد: در پنج دوره نخست، نرخ یادگیری به‌صورت خطی از صفر تا ۴-۱۰×۱ افزایش یافت و پس از آن، با استفاده از راهبرد کاهش کسینوسی (Cosine Decay) به‌تدریج کاهش پیدا کرد. این راهبرد به پایداری فرآیند آموزش و بهبود همگرایی مدل کمک می‌کند.

به‌منظور جلوگیری از بیش‌برازش، از روش Early Stopping با آستانه ۱۰ دوره استفاده شد؛ بدین معنا که اگر معیار خطا (Loss) روی مجموعه اعتبارسنجی طی ۱۰ دوره متوالی بهبود نمی‌یافت، فرآیند آموزش متوقف می‌شد.

نتایج

طریقه تقسیم‌بندی مجموعه‌داده به صورت ۸۰٪ متعلق به مجموعه‌داده آموزشی، ۱۰٪ متعلق به مجموعه‌داده تست و ۱۰٪ متعلق به مجموعه‌داده اعتبارسنجی انجام شده است.

زمانی که خروجی شامل دو کلاس یا بیشتر باشد ماتریس درهم‌ریختگی (Confusion Matrix) ساده‌ترین راه برای اندازه‌گیری کارایی یک مسئله طبقه‌بندی است. صحت(Accuracy)، دقت (Precision)، حساسیت(Sensitivity) و F1-score متداول‌ترین معیارها برای بررسی کارایی الگوریتم‌های طبقه‌بندی است [۲۵] که در این پژوهش نیز از آن‌ها برای گزارش نتایج مدل پیشنهادی استفاده شده است.

همانگونه که در بخش معرفی پایگاه داده تشریح گردید، به علت وجود مسئله نشت داده در پایگاه داده اولیه، در مطالعه حاضر نتایج برای هر دو مجموعه داده LC25000 و LC25000-clean به صورت مقایسه‌ای گزارش می‌شود.

نتایج روی بانک داده LC25000

خلاصه نتایج به دست آمده از تعلیم مدل‌های ResNet50، ResNet152v2 و InceptionV3 در جدول ۱ آمده است. همانطور که مشاهده می‌شود مدل ResNet152v2 بهترین عملکرد را داشته و به نتایج بالایی از دقت، صحت و حساسیت دست یافته است.

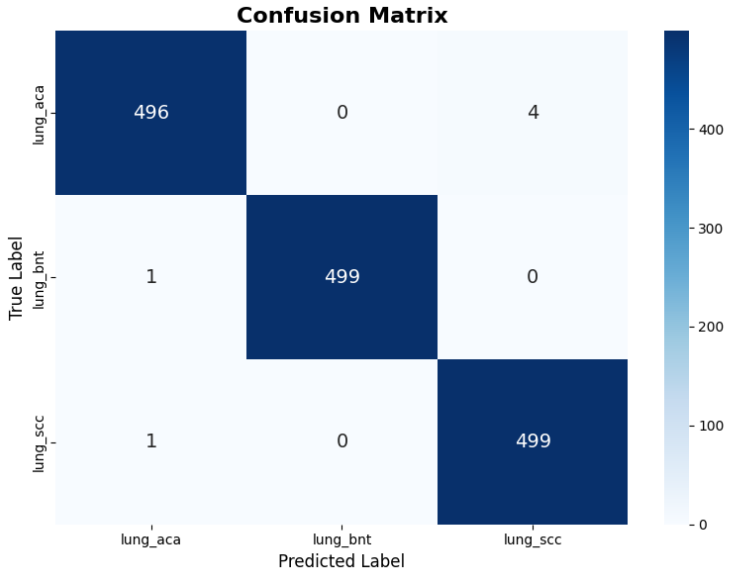
مقاله پژوهشی، صفحات ۱۶۳-۱۷۹
DOI: 10.22034/cmr.2026.8642.3327

در ادامه، جزئیات نتایج مربوط به بهترین شبکه عمیق در قالب ماتریس درهم ریختگی در تصویر ۳ نمایش داده شده است. همانطور که از نتایج واضح است، شبکه در تشخیص طبقه خوش‌خیم بیماری مشکلی نداشته و ۹۹.۸۰٪ موفق عمل می‌کند اما در طبقه‌بندی دو زیررده سرطان ریه، دچار تشخیص اشتباه شده و عملکرد آن کمی افت پیدا می‌کند. تصویر ۴ نیز فرایند تعلیم مدل را در داده‌های آموزش و ارزیابی نشان می‌دهد.

جدول ۲ مقایسه‌ای از نتایج روش پیشنهادی ما با مطالعات قبلی که روی یک بانک داده گزارش شده‌اند، ارائه می‌کند. با توجه به تعداد داده مورد استفاده در تحقیقات گذشته و تفاوت در روش های پیش پردازش تصاویر ، مقایسه مستقیم روش‌های پردازشی منصفانه نیست. اما می توان یک جمع بندی از کلیات روش‌های اعمال شده در پژوهش‌های مختلف و موفقیت نسبی آن‌ها در تشخیص سرطان از روی تصاویر هیستوپاتولوژی داشت. همانطور که در این جدول ملاحظه می‌شود، نتایج این تحقیق نشان می‌دهد که مدل ResNet152v2 صحت بالاتری نسبت پژوهش‌های اخیر داشته است.

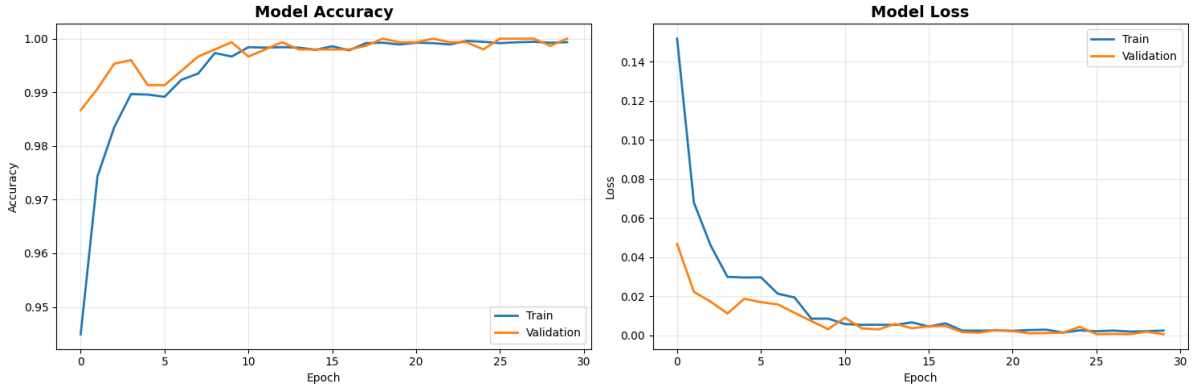
جدول ۱: خلاصه نتایج به دست آمده از تعلیم مدل شبکه عصبی از پیش آموزش دیده

مدل	صحت (٪)	دقت (٪)	حساسیت (٪)	F1-score (٪)
ResNet50	۹۴.۵	۹۵.۳	۹۴.۳	۹۴.۳
ResNet152v2	۹۹.۶	۹۹.۶	۹۹.۶	۹۹.۶
Inception-v3	۹۶.۴	۹۷.۰	۹۶.۳	۹۶.۳



تصویر ۱: نتایج تست شبکه ResNet152v2 در قالب ماتریس درهم ریختگی بر روی بانک داده LC25000

مقاله پژوهشی، صفحات ۱۶۳-۱۷۹
DOI: 10.22034/cmr.2026.8642.3327



تصویر۲: نمودار صحت و زیان تعلیم مدل پس از به کار گیری مدل از پیش آموزش دیده ResNet152v2 روی بانک داده LC25000

جدول ۲: مقایسه نتایج به دست آمده در این پژوهش با پژوهش‌های مشابه بر روی بانک داده LC25000

نتایج برحسب صحت (٪)	مدل استفاده شده	منبع
۹۷.۴	UNet	[۸]
۹۵.۸	Seg-Net	[۹]
۹۳.۰	ResNet50	[۵]
۹۷.۰	CroReLU	[۱۶]
۹۹.۳	feature transformation methods + lightweight CNNs	[۲]
۹۸.۹	ResNet101	[۱]
۹۹.۶	ResNet152v2	پژوهش حاضر

نتایج روی بانک داده LC25000-clean

با توجه به برتری عملکرد شبکه ResNet152v2 در مقایسه با سایر روش‌های به‌کارگرفته‌شده بر روی بانک داده LC25000، در این بخش تمرکز صرفاً بر ارزیابی دقیق این معماری قرار گرفته است. هدف اصلی، تحلیل اثر پدیده‌ی نشت داده و بررسی میزان تأثیر آن بر نتایج نهایی مدل می‌باشد. از آنجا که وجود نمونه‌های مشابه از یک بیمار در مجموعه‌های آموزش و آزمون می‌تواند منجر به برآورد بیش‌ازحد خوش‌بینانه‌ی عملکرد مدل شود، کنترل این مسئله برای دستیابی به ارزیابی واقع‌بینانه ضروری است.

نمونه‌هایی از تصاویر مربوط به کلاس lung_aca که همگی متعلق به یک فرد هستند، در تصویر ۵ نمایش داده شده‌اند. شباهت بالای این تصاویر می‌تواند در صورت تقسیم‌بندی نامناسب داده‌ها، موجب ورود تصاویر بسیار مشابه به مجموعه‌های آموزش و آزمون شود و در نتیجه پدیده نشت داده را رقم بزند. این امر سبب می‌شود مدل به‌جای یادگیری ویژگی‌های تعمیم‌پذیر، الگوهای وابسته به فرد خاص را بیاموزد.

به منظور کنترل این مشکل، فرآیند تفکیک داده‌ها به‌گونه‌ای انجام شد که تصاویر مربوط به هر بیمار صرفاً در یکی از مجموعه‌های آموزش، اعتبارسنجی یا آزمون قرار گیرند. نتایج حاصل از اعمال شبکه بر روی نسخه‌ی کنترل‌شده‌ی بانک داده (LC25000-clean) در جدول ۳ ارائه شده است. همان‌طور که مشاهده می‌شود، با حذف اثر نشت داده، مقادیر دقت نسبت به حالت اولیه اندکی کاهش یافته، اما ارزیابی به‌دست‌آمده واقع‌بینانه‌تر و قابل اتکاتر است.

در ادامه، برای تحلیل دقیق‌تر عملکرد مدل، ماتریس درهم‌ریختگی بهترین شبکه عمیق در تصویر ۶ نمایش

DOI: 10.22034/cmr.2026.8642.3327
مقاله پژوهشی، صفحات ۱۶۳-۱۷۹

داده شده است. بررسی این ماتریس نشان می‌دهد که مدل در تشخیص نمونه‌های خوش‌خیم عملکرد بسیار مطلوبی داشته و با دقتی در حدود ۹۹.۸ درصد این طبقه را به‌درستی شناسایی کرده است. با این حال، در تفکیک دو زیررده سرطان ریه (به‌ویژه میان کلاس‌های بدخیم با الگوهای بافتی مشابه)، میزان خطای طبقه‌بندی افزایش یافته و بخشی از نمونه‌ها به‌صورت متقابل اشتباه تشخیص داده شده‌اند. این موضوع بیانگر شباهت ویژگی‌های بافت‌شناختی این دو زیررده و چالش ذاتی در تمایز آن‌ها است.

پس از رفع مشکل نشت داده در مجموعه‌داده LC25000 و ارزیابی مدل بر روی LC25000-clean، شاهد کاهش نسبی در شاخص‌های عملکردی بودیم. صحت مدل از مقادیر نزدیک به ۹۹.۶٪ (در حالت نشت داده) به ۹۶.۹٪ کاهش یافت. این افت عملکرد، که عمدتاً ناشی از کاهش حساسیت در کلاس lung_aca (با ۸.۸۱٪ خطا) و کاهش دقت در کلاس lung_scc (با ۸.۷۳٪ خطا) است، نشان‌دهنده حذف اثرات مخرب نشت داده و دستیابی به ارزیابی واقع‌بینانه‌تر از قابلیت تعمیم‌پذیری مدل می‌باشد. با وجود این کاهش، صحت ۹۶.۹٪ عملکرد بسیار مطلوبی برای غربالگری تصاویر هیستوپاتولوژی محسوب می‌شود.

به طور کلی، نتایج نشان می‌دهد که هرچند شبکه ResNet152v2 توانایی بالایی در استخراج ویژگی‌های عمیق و تشخیص الگوهای پاتولوژیک دارد، اما کنترل نشت داده نقش تعیین‌کننده‌ای در ارائه ارزیابی صحیح از عملکرد مدل ایفا می‌کند و باید به‌عنوان یک مرحله‌ی اساسی در مطالعات مبتنی بر داده‌های پزشکی مدنظر قرار گیرد.

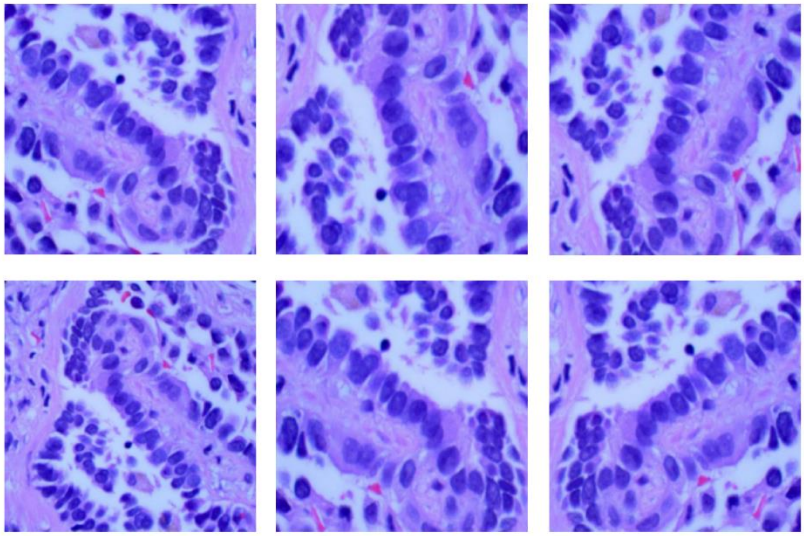
به منظور درک بهتر ویژگی‌های مورفولوژیکی که مدل برای طبقه‌بندی تصاویر هیستوپاتولوژی استفاده می‌کنند، از تکنیک Grad-CAM (Gradient-weighted Class Activation Mapping) استفاده شد [۲۲]. این روش با محاسبه گرادیان‌های خروجی نسبت به آخرین لایه کانولوشنال، نقشه حرارتی ایجاد می‌کند که نواحی تأثیرگذار در تصمیم‌گیری مدل را مشخص می‌سازد.

در این پژوهش، برای هر تصویر، نقشه حرارتی Grad-CAM محاسبه و بر روی تصویر اصلی منطبق گردید. نواحی با فعال‌سازی قوی (رنگ‌های گرم) نشان‌دهنده بخش‌هایی از تصویر هستند که بیشترین تأثیر را در تصمیم‌گیری مدل داشته‌اند.

تصویر ۷ نقشه‌های حرارتی Grad-CAM را برای ۵ نمونه تصادفی از مجموعه آزمون نشان می‌دهد. این نمونه‌ها شامل سه نمونه کارسینوم سلول سنگفرشی (lung_scc)، یک نمونه بافت سالم ریه (lung_n) و یک نمونه آدنوکارسینوما (lung_aca) هستند که همگی توسط مدل ResNet152v2 به درستی طبقه‌بندی شده‌اند. نقشه‌های حرارتی در این تصویر نشان می‌دهند که مدل تمرکز خود را بر روی نواحی بافتی مرتبط با ویژگی‌های پاتولوژیک هر کلاس سرطان ریه قرار داده است؛ برای مثال، در نمونه‌های مربوط به سلول‌های سنگفرشی، توجه مدل عمدتاً روی مناطق متراکم سلولی و هسته‌های تیره‌تر متمرکز است، در حالی که در نمونه‌های آدنوکارسینوم، نواحی فعال به شکل متمرکزتر و مرتبط با ساختارهای غده‌ای دیده می‌شود. همچنین در نمونه‌های بافت نرمال، مدل بیشتر روی ساختارهای منظم بافتی تمرکز کرده است. تصاویر همپوشان نیز تأیید می‌کنند که نواحی فعال مدل دقیقاً بر روی بخش‌های کلیدی تصویر قرار دارند و توجه پراکنده یا غیر مرتبط مشاهده نمی‌شود. این رفتار مدل نشان‌دهنده یادگیری ویژگی‌های معنی‌دار و مرتبط با بافت‌شناسی سرطان است و از توجه به نویز یا پس‌زمینه جلوگیری شده است. به طور کلی، نتایج Grad-CAM

مقاله پژوهشی، صفحات ۱۶۳-۱۷۹
DOI: 10.22034/cmr.2026.8642.3327

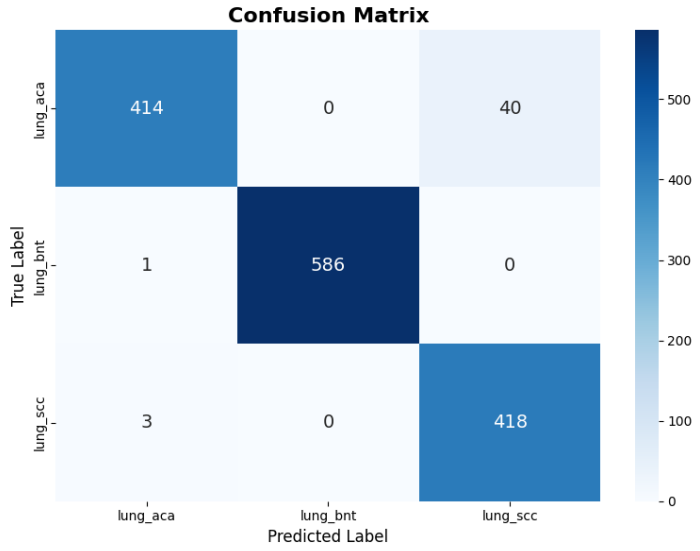
نشان می‌دهد که مدل نه تنها قادر به تشخیص دقیق تیپ‌های مختلف سرطان ریه است، بلکه از نظر قابلیت تبیین‌پذیری نیز عملکرد قابل قبولی دارد و بر نواحی مرتبط و پاتولوژیک تمرکز می‌کند که این امر اهمیت بالایی در کاربردهای پزشکی دارد.



تصویر ۵: نمونه تصاویر مربوط به کلاس lung_aca که همگی متعلق به یک فرد می‌باشند.

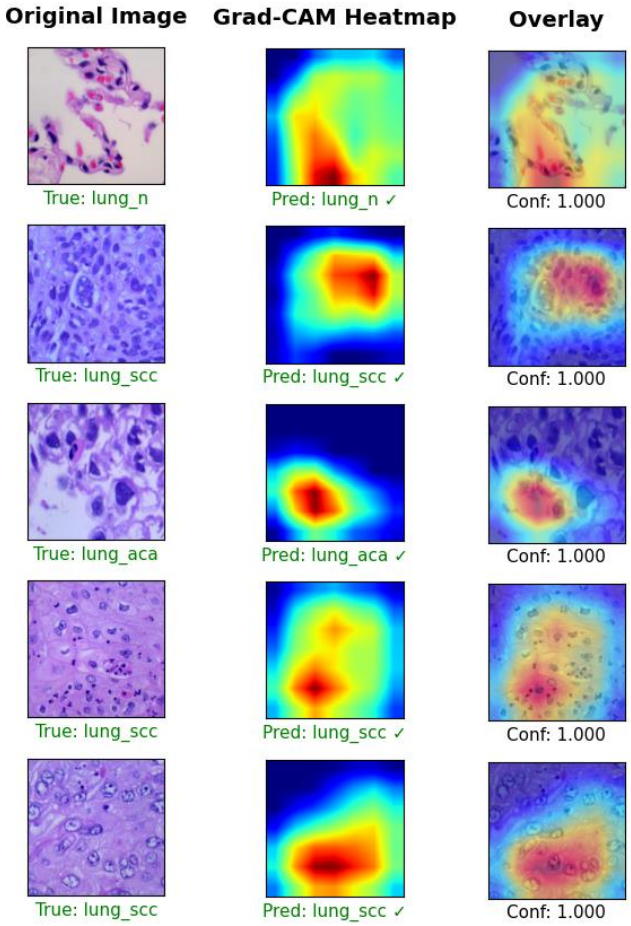
جدول ۳: خلاصه نتایج به دست آمده از تعلیم مدل ResNet152v2 از پیش آموزش دیده بر روی بانک داده LC25000-clean.

مدل	صحت (٪)	دقت (٪)	حساسیت (٪)	F1-score (٪)
ResNet152v2	۹۶.۹	۹۶.۷	۹۶.۷	۹۶.۶



تصویر ۶: نتایج تست شبکه ResNet152v2 در قالب ماتریس درهم ریختگی بر روی بانک داده LC25000-clean.

مقاله پژوهشی، صفحات ۱۶۳-۱۷۹
DOI: 10.22034/cmr.2026.8642.3327



تصویر ۷: نمایش نتایج آزمون Grad-CAM روی تعدادی تصاویر تست بانک داده LC25000-clean. در ستون اول تصاویر اصلی، در ستون دوم نقشه‌های حرارتی Grad-CAM و در ستون سوم تصاویر همپوشان با نقشه‌های توجه نمایش داده شده‌اند. مقدار اطمینان (Conf) مدل در پیش‌بینی هر نمونه در زیر تصاویر همپوشان آورده شده است که نشان‌دهنده اعتماد بالای مدل به تصمیمات خود می‌باشد.

بحث

در این مطالعه، تصاویر هیستوپاتولوژی بافت ریه با استفاده از روش‌های یادگیری عمیق برای تشخیص و طبقه‌بندی سرطان ریه به سه گروه خوش‌خیم، ادنوکارسینوم و کارسینومای سلول سنگفرشی مورد بررسی قرار گرفت. یکی از چالش‌های اصلی، حجم محدود داده‌ها بود که آموزش مستقیم شبکه‌های عصبی عمیق را دشوار می‌کرد. برای مقابله با این محدودیت، از یادگیری انتقالی بهره گرفته شد که امکان بهره‌گیری از شبکه‌های پیش‌آموخته و انتقال دانش از مجموعه‌های بزرگ‌تر را فراهم کرد.

نتایج حاصل نشان‌دهنده توانایی روش‌های یادگیری عمیق در تفسیر تصاویر هیستوپاتولوژی و شناسایی ویژگی‌های بافتی مرتبط با سرطان ریه بود. در میان مدل‌های ارزیابی شده، شبکه ResNet152v2 بهترین عملکرد کلی را ارائه داد و توانست دقت بسیار بالایی در تشخیص تصاویر خوش‌خیم و کارسینومای سلول سنگفرشی (حدود ۹۹.۹۸٪) به دست آورد. با این حال، در طبقه‌بندی ادنوکارسینوم دقت مدل کمی کاهش یافته و به ۹۹.۲۰٪ رسید. این اختلاف عملکرد نشان‌دهنده نقاط ضعفی است که با تنظیم معماری شبکه، افزودن دانش بالینی، یا جمع‌آوری داده‌های بیشتر قابل بهبود است. بررسی سایر مدل‌های پیش‌آموخته یا تنظیم دقیق مدل فعلی نیز می‌تواند منجر به ارتقای قابلیت تفکیک زیررده‌های خاص شود.

مقاله پژوهشی، صفحات ۱۶۳-۱۷۹
DOI: 10.22034/cmr.2026.8642.3327

با اصلاح نشت داده در مجموعه LC25000 و ارزیابی مجدد مدل بر روی نسخه LC25000-clean، کاهش نسبی در شاخص‌های عملکرد مشاهده شد. صحت ResNet152v2 از حدود ٪۹۹.۶ (با نشت داده) به ٪۹۶.۹ کاهش یافت. این افت عمدتاً ناشی از کاهش حساسیت در کلاس lung_aca (٪۸.۸۱ خطا) و کاهش دقت در کلاس lung_scc (٪۸.۷۳ خطا) بود. این کاهش عملکرد، در واقع نمایانگر ارزیابی واقعی‌تر و عادلانه‌تر قابلیت تعمیم‌پذیری مدل است و با وجود آن، دقت ٪۹۶.۹ همچنان برای کاربردهای غربالگری تصاویر هیستوپاتولوژی عملکردی بسیار مطلوب محسوب می‌شود.

برای غلبه بر چالش ماهیت غیرشفاف مدل‌ها، تحلیل Grad-CAM نشان داد که مدل بر ویژگی‌های مورفولوژیک معنادار و همخوان با معیارهای پاتولوژیست‌ها تمرکز دارد. این موضوع از دو جهت حائز اهمیت است: اولاً اعتبار مدل را تأیید می‌کند (زیرا به‌جای همبستگی‌های سطحی مانند رنگ‌آمیزی، به شواهد بافتی واقعی توجه دارد) و ثانیاً نقشه‌های حرارتی حاصل، اعتماد پاتولوژیست‌ها را افزایش داده و کاربردهای آموزشی و شناسایی موارد مرزی را تسهیل می‌کنند.

بااین‌حال، مطالعه با محدودیت‌های اساسی نظیر عدم اعتبارسنجی بالینی مستقل، استفاده از تصاویر قطعه‌شده به‌جای اسلاید کامل، تک‌مرکزی بودن داده‌ها، و عدم تطابق با پیچیدگی دیتاست‌های واقعی نظیر TCGA [۱۸] مواجه است. در نتیجه، نتایج فعلی بیشتر نمایانگر عملکرد در محیط آزمایشگاهی بوده و نمی‌توانند به‌عنوان معیار مستقیم کارآیی بالینی در نظر گرفته شوند.

بر همین اساس، پژوهش‌های آتی باید بر اعتبارسنجی روی دیتاست‌های مستقل بالینی مانند TCGA و CPTAC یا داده‌های چندمرکزی [۱۸]-[۱۱]، ارزیابی روی اسلایدهای کامل، انجام مطالعات چندمرکزی با در نظر گرفتن تنوع پروتکل‌ها و جمعیت‌شناسی، و نیز مقایسه مستقیم عملکرد مدل با تشخیص پاتولوژیست‌ها متمرکز شوند.

نتیجه گیری

یافته‌های این پژوهش بر اهمیت ارزیابی مداوم و بهبود مدل‌های یادگیری ماشین در تشخیص‌های پزشکی تأکید دارند. اگرچه روش‌های یادگیری عمیق پتانسیل بالایی در تفسیر تصاویر هیستوپاتولوژی و تشخیص سرطان ریه نشان داده‌اند، اما دستیابی به کاربرد بالینی پایدار مستلزم عبور از مرحله آزمایشگاهی و انجام اعتبارسنجی‌های دقیق بر روی داده‌های مستقل و متنوع است.

تضاد منافع

نویسندگان اعلام می کنند که هیچ تضاد منافعی ندارند.

سپاسگزاری

این اثر تحت حمایت مادی صندوق حمایت از پژوهشگران و فناوران کشور (INSF) برگرفته از طرح شماره ۴۰۰۲۳۸۹ انجام شده است. برای اصلاح بخش‌هایی از متن، از ابزارهای هوش مصنوعی استفاده شده است.

فهرست منابع

1. Abd El-Ghany, S., Azad, M., and Elmogy, M., “Robustness fine-tuning deep learning model for cancers diagnosis based on histopathology image analysis,” *Diagnostics*, vol. 13, no. 4, 2023.
2. Attallah, O., Aslan, M. F., and Sabanci, K., “A framework for lung and colon cancer diagnosis via

مقاله پژوهشی، صفحات ۱۶۳-۱۷۹
DOI: 10.22034/cmr.2026.8642.3327

lightweight deep learning models and transformation methods," *Diagnostics*, vol. 12, no. 12, 2022.

3. Batchkala, G., Li, B., and Rittscher, J., "Evaluating histopathology foundation models for few-shot tissue clustering: An application to LC25000 augmented dataset cleaning," in *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Springer, 2025.
4. Borkowski, A. A., Bui, M. M., Thomas, L. B., and Wilson, C. P., "Lung and colon cancer histopathological image dataset (LC25000)," *arXiv preprint arXiv:1912.12142*, 2019.
5. Bukhari, S. U. K., Syed, A., Bokhari, S. K. A., Hussain, S. S., Armaghan, S. U., and Shah, S. S. H., "The histological diagnosis of colonic adenocarcinoma by applying partial self supervised learning," *medRxiv*, 2020.
6. Chen, C., et al., "An annotation-free whole-slide training approach to pathological classification of lung cancer types using deep learning," *Nature Communications*, vol. 12, no. 1193, 2021. doi: 10.1038/s41467-021-21467-y.
7. Coudray, N., et al., "Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning," *Nature Medicine*, vol. 24, pp. 1559–1567, 2018.
8. Dabass, M. and Dabass, J., "An Atrous convolved hybrid Seg-Net model with residual and attention mechanism for gland detection and segmentation in histopathological images," *Computers in Biology and Medicine*, vol. 155, 2023.
9. Dabass, M., Dabass, J., Vashisth, S., and Vig, R., "A hybrid U-Net model with attention and advanced convolutional learning modules for simultaneous gland segmentation and cancer grade prediction in colorectal histopathological images," *Intelligence-Based Medicine*, vol. 7, 2023.
10. Das, A., Nair, M. S., and Peter, S. D., "Computer-aided histopathological image analysis techniques for automated nuclear atypia scoring of breast cancer: A review," *Journal of Digital Imaging*, vol. 33, no. 2, pp. 437–456, 2020. doi: 10.1007/s10278-019-00295-z.
11. Ellis, M. J., et al., "Connecting genomic alterations to cancer biology with proteomics: The NCI Clinical Proteomic Tumor Analysis Consortium," *Cancer Discovery*, vol. 3, no. 10, pp. 1108–1112, 2013.
12. Hashemzadeh, H., Shojaeilangari, S., and Allahverdi, A., "A combined microfluidic deep learning approach for lung cancer cell high throughput screening toward automatic cancer screening applications," *Scientific Reports*, vol. 11, no. 1, pp. 1–10, 2021. doi: 10.1038/s41598-021-89352-8.
13. Hassanpour, S. H. and Dehghani, M., "Review of cancer from perspective of molecular," *Journal of Cancer Research and Practice*, vol. 4, no. 4, 2017. doi: 10.1016/j.jcrpr.2017.07.001.
14. He, K., Zhang, X., Ren, S., and Sun, J., "Deep residual learning for image recognition," in *Proc. IEEE*

DOI: 10.22034/cmr.2026.8642.3327 مقاله پژوهشی، صفحات ۱۶۳-۱۷۹

Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770–778.

15. Kanavati, F., et al., "Weakly-supervised learning for lung carcinoma classification using deep learning," *Scientific Reports*, vol. 10, no. 1, pp. 1–11, 2020. doi: 10.1038/s41598-020-66333-x.
16. Liu, Y., et al., "CroReLU: Cross-crossing space-based visual activation function for lung cancer pathology image recognition," *Cancers*, vol. 14, no. 21, 2022.
17. Lu, M. Y., et al., "Data-efficient and weakly supervised computational pathology on whole-slide images," *Nature Biomedical Engineering*, vol. 5, no. 6, pp. 555–570, 2021. doi: 10.1038/s41551-020-00682-w.
18. National Cancer Institute, "The Cancer Genome Atlas (TCGA)," [Online]. Available: https://www.cancer.gov/ccg/research/genome-sequencing/tcga.
19. Nayak, T., et al., "Automated histopathological detection and classification of lung cancer with an image pre-processing pipeline and spatial attention with deep neural networks," *Cogent Engineering*, vol. 11, no. 1, 2024.
20. Parra-Medina, R., et al., "Deep learning in histopathology images for prediction of oncogenic driver molecular alterations in lung cancer: A systematic review and meta-analysis," *Translational Lung Cancer Research*, vol. 14, no. 5, pp. 1756–1769, May 2025. doi: 10.21037/tlcr-2024-1196.
21. Roche, J. J., Seyedshahi, F., Rakovic, K., Thu, A. W., Le Quesne, J., and Blyth, K. G., "Current and future applications of artificial intelligence in lung cancer and mesothelioma," *Thorax*, vol. 80, no. 12, pp. 957–965, Nov. 2025. doi: 10.1136/thorax-2024-222054.
22. Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D., "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, 2017, pp. 618–626. doi: 10.1109/ICCV.2017.74.
23. Shmatko, A., Ghaffari Laleh, N., Gerstung, M., and Kather, J. N., "Artificial intelligence in histopathology: Enhancing cancer research and clinical oncology," *Nature Cancer*, vol. 3, pp. 1026–1038, 2022.
24. Siegel, R. L., Giaquinto, A. N., and Jemal, A., "Cancer statistics, 2024," *CA: A Cancer Journal for Clinicians*, vol. 74, pp. 12–49, 2024.
25. Sokolova, M. and Lapalme, G., "A systematic analysis of performance measures for classification tasks," *Information Processing & Management*, vol. 45, no. 4, pp. 427–437, 2009. doi: 10.1016/j.ipm.2009.03.002.
26. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z., "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 2818–2826.

مقاله پژوهشی، صفحات ۱۶۳-۱۷۹
DOI: 10.22034/cmr.2026.8642.3327

27. Teramoto, A., Tsukamoto, T., Kiriyama, Y., and Fujita, H., "Automated classification of lung cancer types from cytological images using deep convolutional neural networks," *BioMed Research International*, vol. 2017, 2017. doi: 10.1155/2017/4067832.
28. Wei, J. W., Tafe, L. J., Linnik, Y. A., Vaickus, L. J., Tomita, N., and Hassanpour, S., "Pathologist-level classification of histologic patterns on resected lung adenocarcinoma slides with deep neural networks," *Scientific Reports*, vol. 9, no. 1, pp. 1–8, 2019. doi: 10.1038/s41598-019-40041-7.

مقاله پژوهشی، صفحات ۱۶۳-۱۷۹
DOI: 10.22034/cmr.2026.8642.3327

Automated Classification of Lung Cancer from Histopathological Images using Deep Learning

Seyedehsamaneh Shojaeilangari[1]*, Abolfazl Bazazpour [2]

[1] Biomedical Engineering Group, Department of Electrical and Information Technology, Iranian Research Organization for Science and Technology (IROST), 33535111, Tehran, Iran

[2] GIS Group, Faculty of Remote Sensing and GIS, Shahid Beheshti University, Tehran, Iran

* Corresponding author: E-mail: s.shojaie@irost.ir

Abstract:

Cancer is one of the leading causes of death worldwide, and its early detection plays a key role in effective management. The standard diagnostic method is biopsy and manual examination of histopathological slides by pathologists, which is time-consuming, requires high expertise, and is prone to error. Therefore, the use of computer-aided diagnostic systems based on artificial intelligence can help improve the efficiency of medical image interpretation. AI methods have multiplied the ability to extract quantitative information from digital histopathological images and are expected to reduce the workload for human experts while improving the accuracy and consistency of histopathological reports. In this research, various deep learning models, including ResNet50, ResNet152v2, and Inception-v3, were employed for lung cancer detection and subtype classification from histopathological images. The results of this study on the LC25000 dataset, utilizing transfer learning, demonstrate the significant potential of the proposed approach in diagnosing this disease. Among these, the ResNet152v2 network achieved the best performance, attaining 99.6% accuracy, 99.6% precision, and 99.6% sensitivity. However, after applying data leakage control and rigorous patient-wise separation of images into training and test sets, these values decreased to 96.9%, 96.7%, and 96.7%, respectively, providing a more realistic evaluation of the model. Overall, the findings indicate the high capacity of deep learning methods in classifying lung cancer subtypes. Nevertheless, the generalizability of the results to clinical settings requires validation on independent and diverse datasets, including whole-slide images, in future studies.

Keywords : Histopathological Images, Lung Cancer, Automated Diagnostic Systems, Convolutional Neural Networks, Deep Learning