

دریافت مقاله: ۱۴۰۳/۱۱/۰۶

پذیرش مقاله: ۱۴۰۳/۱۲/۱۸

نوع مقاله: پژوهشی

تشخیص هیستوپاتولوژی سرطان پستان توسط شبکه‌های عصبی هم‌آمیختی بهبودیافته با سازوکارهای توجه

فرید کریمی

گروه پاتوبیولوژی، دانشکده دامپزشکی، دانشگاه شهرکرد، شهرکرد، ایران.

پست الکترونیک: farid8kd35@gmail.com

مریم کریمی*

گروه علوم کامپیوتر، دانشکده علوم ریاضی، دانشگاه شهرکرد، شهرکرد، ایران.

پست الکترونیک: ma.karimi@sku.ac.ir

چکیده

این تحقیق نه تنها برتری سازوکارهای توجه را تأیید می‌کند، بلکه راه‌حلی عملی برای کاهش خطاهای تشخیصی و حمایت از تصمیم‌گیری‌های بالینی ارائه می‌دهد. نتایج این مطالعه می‌تواند به توسعه ابزارهای تشخیصی پیشرفته و بهبود نتایج درمانی بیماران کمک کند.

واژه‌های کلیدی: سرطان پستان، تشخیص هیستوپاتولوژیک، شبکه‌های عصبی هم‌آمیختی، سازوکار توجه، DANet, PANet, CBAM.

مقدمه

سرطان پستان به عنوان یک چالش مهم در بهداشت عمومی مطرح می‌باشد و این امر با افزایش نرخ بروز آن برجسته‌تر می‌شود. مطالعات نشان می‌دهند این نرخ سالانه ۱ درصد افزایش می‌یابد و این مسئله با وجود پیشرفت‌های صورت گرفته در زمینه درمان، با افزایش نامتناسب سرطان پستان در برخی از زیرجمعیت‌ها پیچیده‌تر می‌شود، به طوری که این افزایش در زنان زیر

سرطان پستان به عنوان یکی از چالش‌های مهم بهداشت عمومی، نیازمند تشخیص دقیق و به موقع است. این مطالعه به بررسی تأثیر سازوکارهای توجه در شبکه‌های عصبی هم‌آمیختی (CNN) برای تحلیل تصاویر هیستوپاتولوژیک سرطان پستان می‌پردازد. با استفاده از سه سازوکار توجه شامل DANet, PANet و CBAM، یک مدل CNN سفارشی طراحی شده است که به طور سیستماتیک عملکرد این سازوکارها را ارزیابی می‌کند. نتایج نشان می‌دهد که DANet با دقت ۹۲/۰۲ درصد و AUC-ROC برابر با ۹۷/۷۰ درصد، بهترین عملکرد را در مقایسه با CBAM و PANet دارد. این برتری به توانایی DANet در مدل‌سازی هم‌زمان وابستگی‌های مکانی و کانالی مربوط می‌شود که به آن اجازه می‌دهد بر مناطق کلیدی تصاویر تمرکز کند. همچنین، CBAM با دقت قابل توجهی در شناسایی ویژگی‌های مهم، نتایج خوبی ارائه می‌دهد، در حالی که PANet به دلیل ساختار سلسله‌مراتبی خود، عملکرد ضعیف‌تری داشت.

* نویسنده مسئول

سازوکارهای توجه با تقلید از فرآیند توجه انسان شبکه را به نحوی هدایت می‌کنند تا بر مناطق مهم تصویر تمرکز کند و بازنمایی ویژگی‌های مرتبط را تقویت کند [۵].

مطالعات اخیر نشان می‌دهند تأثیر سازوکارهای توجه در تحلیل تصاویر هیستوپاتولوژیک بسیار زیاد است. به عنوان مثال، شبکه DMBANet که توسط Nagi و همکاران در سال ۲۰۲۴ ارائه شد، با ادغام سازوکارهای توجه کانالی و فضایی در یک معماری چندشاخه CNN، به دقت ۹۸ درصد در مجموعه داده PathoIMG دست یافت [۶]. همچنین، SANAS-Net که توسط Souza و همکاران معرفی شد، از توجه چندسر^۷ در بلوک‌های ترنسفورمر برای شناسایی روابط فضایی متنوع استفاده کرد و به دقت ۸۹/۹۵ درصد در تشخیص موارد بدخیم از خوش خیم سرطان پستان در تصاویر ماموگرام^۸ رسید [۷]. علاوه بر این، ادغام سازوکارهای توجه مانند CBAM و سازوکار توجه غیرمحلی^۹ با معماری‌های پیشرفته‌ای مانند ResNet و DenseNet، دقت و تفسیرپذیری مدل‌ها را در بخش‌بندی تومورهای پستان بهبود بخشیده است. مطالعات ذکر شده پتانسیل سازوکارهای توجه را در ارتقای عملکرد مدل‌های CNN برای تشخیص سرطان پستان تأیید می‌کنند [۸].

این مطالعه با بررسی مجزای سه سازوکار توجه شامل DANet، PANet و CBAM در یک مدل CNN سفارشی، رویکردی نوین برای تشخیص هیستوپاتولوژیک سرطان پستان ارائه می‌دهد. ما عملکرد این سازوکارها را به طور سیستماتیک ارزیابی کرده‌ایم و نشان داده‌ایم که DANet با دقت ۹۲/۰۲ درصد و AUC-ROC برابر با ۹۷/۷۰ درصد، بهترین عملکرد را داشته و از CBAM و PANet پیشی گرفته است. معماری CNN سفارشی ما، که برای یادگیری ویژگی‌های چندمقیاسی بهینه‌سازی شده، هم جزئیات هسته‌ای و هم سازمان‌بندی بافت را شناسایی می‌کند که برای تشخیص دقیق سرطان حیاتی است. با ترکیب

۵۰ سال با شدت بیشتری رخ می‌دهد، در نتیجه تشخیص دقیق و زودهنگام آن اهمیت به‌سزایی در بهبود نتایج درمان بیماران دارد [۱].

تشخیص دقیق و به موقع تومورهای پستان که پس از بیوپسی تومور از طریق روش هیستوپاتولوژی به عنوان ابزار اصلی تشخیص انجام می‌شود، در بهبود نتایج درمان بیماران بسیار حائز اهمیت است؛ این روش شامل به دست آوردن نمونه‌های بافتی از مناطق مشکوک با استفاده از روش‌های مختلف بیوپسی نظیر آسپیراسیون^۱ با سوزن ظریف، بیوپسی^۲ با سوزن مرکزی، بیوپسی با کمک خا^۳ و بیوپسی جراحی است [۲]. این نمونه‌ها بر روی لام‌های میکروسکوپی قرار داده می‌شوند و پس از آن با استفاده از هماتوکسیلین و ائوزین (H&E)^۴ رنگ‌آمیزی می‌شوند تا وضوح ساختارهای بافتی و اجزای سلولی برای تجزیه و تحلیل بهبود یابد، تصاویری که از روی این لام‌ها تهیه می‌شوند تصویر هیستوپاتولوژیک^۵ نامیده می‌شوند [۳]. با این حال، تحلیل این تصاویر توسط پزشک، به دلیل پیچیدگی الگوهای مربوط به رشد سلول‌ها و تنوع ذاتی آنها، زمان‌بر بوده و مستعد خطاهای انسانی است. این امر ضرورت توسعه سیستم‌های تشخیص کمکی کامپیوتری (CAD) را به عنوان ابزار کمکی برای افزایش سرعت، دقت و کاهش هزینه‌های تشخیصی برجسته می‌سازد.

در سال‌های اخیر، یادگیری عمیق، به ویژه شبکه‌های عصبی هم‌آمیختی (CNN)^۶، به عنوان رویکردی قدرتمند در تحلیل تصاویر پزشکی ظاهر شده‌اند [۴]. با این حال، چالش اصلی در استفاده از CNN‌ها برای تصاویر هیستوپاتولوژیک، استخراج ویژگی‌های متمایزکننده و مرتبط با بیماری از این تصاویر پیچیده است. به همین دلیل، محققان به دنبال راهکارهایی برای بهبود عملکرد CNN‌ها از طریق ادغام سازوکارهای توجه^۶ بوده‌اند.

1-Aspiration

2-Biopsy

3-Hematoxylin & Eosin

4-Histopathologic

5-Convolutional Neural Networks

6-Attention Mechanisms

7-Multi-head Attention

8-Mamogram

9-Non-Local Attention

عملکرد را با مدل DenseNet201 به دست آوردند که دقت ۹۱/۳۷ درصد را در بزرگنمایی ۲۰۰ برابر به دست آورد [۱۰].

سازوکارهای توجه با وزن دهی به ویژگی های مهم، دقت وظایفی مانند طبقه بندی، تقسیم بندی و تشخیص را افزایش می دهند و به تشخیص بهتر و درک الگوهای پیچیده تر در تصاویر می توانند کمک کنند. نویسندگان در [۱۱] شبکه ای به نام PANet پیشنهاد دادند که از دیدگاه یادگیری متریک به این مشکل نگاه می کنند. PANet نمایه های اولیه ای مخصوص هر کلاس را از تصاویر حاشیه نویسی شده پشتیبان در یک فضای تعبیه یاد می گیرد و سپس بخش بندی را با تطبیق هر پیکسل از تصویر پرسش با این نمایه ها انجام می دهد. با استفاده از یادگیری متریک غیرپارامتری^{۱۰}، PANet نمایه های باکیفیت، نمایانگر و تمایز بخشی برای هر کلاس را ارائه می دهد. علاوه بر این، این شبکه تنظیم منظم سازی هم تراز می نمایه ها بین تصاویر پشتیبان و پرسش را معرفی می کند تا از دانش موجود بهتر استفاده کرده و قابلیت تعمیم بهتری در بخش بندی با نمونه های کم داشته باشد.

فو و همکاران^{۱۱} DANet را پیشنهاد کردند که از سازوکار توجه خودکار برای استخراج وابستگی های متنی غنی استفاده می کند. برخلاف روش های سنتی که به تلفیق ویژگی های چندمقیاسی متکی هستند، آن ها DANet را طراحی کردند تا به صورت تطبیقی ویژگی های محلی را با وابستگی های جهانی ترکیب کند. شبکه آن ها شامل دو ماژول توجه است؛ ماژول توجه موقعیت که وابستگی های مکانی را با جمع وزنی ویژگی های تمام موقعیت ها مدل سازی می کند و ماژول توجه کانال که با ترکیب ویژگی های وابسته میان تمامی کانال ها، نقشه های کانال مرتبط را مشخص می کند. با جمع کردن خروجی های این دو ماژول، آن ها نمایه سازی ویژگی ها را بهبود داده و به نتایج دقیق تری در بخش بندی دست یافتند [۱۲].

روش های پیشرفته افزایش داده^{۱۲}، تنظیم ابرپارامترها^{۱۳} و سازوکارهای توجه، قابلیت اطمینان و تفسیرپذیری مدل را بهبود بخشیده ایم. این پژوهش پلی بین پیشرفت های نظری و کاربردهای عملی ایجاد کرده و ابزاری قابل اعتماد برای بهبود تشخیص سرطان پستان و پشتیبانی از تصمیم گیری های بالینی ارائه می دهد.

۲. پژوهش های مرتبط

با مشخص شدن پتانسیل بالای شبکه های عصبی پیچشی در تحلیل تصاویر مختلف، آرایو خو و همکاران به بررسی یک مدل CNN برای طبقه بندی تصاویر حاصل از بیوپسی^{۱۴} تومور پستان که با روش E&H رنگ آمیزی شدند و در چهار کلاس (بافت نرمال، ضایعه خوش خیم، کارسینوم درجا و کارسینوم تهاجمی^{۱۵}) قرار گرفتند و همچنین یک طبقه بندی دودویی کارسینوم/ غیرکارسینوم پرداختند. معماری پیشنهادی CNN را برای ثبت اطلاعات چندمقیاسی، شامل جزئیات هسته ای و سازماندهی کلی بافت، طراحی کردند و از این رو امکان استفاده از آن برای تصاویر هیستوپاتولوژی تمام اسلاید فراهم شد. علاوه بر این، ویژگی های استخراج شده توسط CNN برای آموزش یک طبقه بندی کننده ماشین بردار پشتیبان (SVM) نیز استفاده شد. نتایج گزارش شده نشان دهنده دقت ۷۷/۸ درصد برای طبقه بندی چهار کلاسه و ۸۳/۳ درصد برای طبقه بندی دودویی برای موارد سرطان بود [۹]. سپس سیمونیان و همکاران، یک مجموعه داده ترکیبی از دو مجموعه داده محبوب BreakHis و Histo ایجاد کردند و تصاویر را به دو دسته بدخیم و خوش خیم تقسیم کردند. در این مطالعه از یادگیری انتقالی^{۱۶} با چهار مدل پیش آموزش دیده شامل DenseNet201، ResNet50، ResNet101 و MobileNet-v2 استفاده کردند و بهترین

10-Data Augmentation

11-Hyperparameters

12-Biopsy

13-Normal tissue, Benign lesion, Carcinoma in situ, Invasive carcinoma

14-Transfer Learning

15-Nonparametric Metric Learning

16-Dual Attention Network

دقت ۹۲/۴۵ درصد، حساسیت ۰/۹۸، دقت ۰/۹۱ و امتیاز F ۰/۹۴، از مدل‌های دیگری مانند CNN, AlexNet, ResNet34 و Feed-Forward Neural Network بهتر عمل کرده است [۱۵].

در [۱۶]، پژوهشگران یک معماری جدید CNN برای طبقه‌بندی سرطان پستان خوش‌خیم و بدخیم در تصاویر هیستوپاتولوژیک معرفی کردند. برای بهبود ارائه و استفاده از اطلاعات ویژگی‌ها، از DenseNet به عنوان بلوک ساختاری اصلی استفاده و آن را با ماژول فشار و برانگیختگی (SENet)^{۱۹} ترکیب کردند، در ادامه پژوهشگران یک روش برای ادغام مکانیزم توجه SENet در معماری DenseNet121 به نام IDSnet معرفی کردند که این روش همچنین با استفاده از تلفیق میانگین سراسری^{۲۰}، که پیچیدگی مدل را کاهش می‌دهد، به محدودیت‌های منابع محاسباتی و مشکل پیش‌پردازش می‌پردازد. مدل IDSnet با استفاده از مدل پیش‌آموزش دیده و تنظیم شده DenseNet-121 ویژگی‌ها را از تصاویر هیستوپاتولوژیک استخراج کرده و با استفاده از ماژول‌های SENet این ویژگی‌ها را اصلاح می‌کند تا اطلاعات جهانی ضروری را به طور مؤثری استخراج کند. نتایج ارزیابی نشان می‌دهد عملکرد IDSnet در مقایسه با مدل‌های VGG-16 و ResNet50 عملکرد بهتری دارد و دقت شناسایی تصویر بین ۸۴/۵ درصد تا ۸۹/۱ درصد دارد [۱۶].

در سال ۲۰۲۴ پژوهشگران در مطالعه‌ای با ادغام سازوکارهای توجه در معماری MobileNetV3، توانایی مدل در تمرکز بر ویژگی‌های کلیدی را بهبود بخشیدند به این گونه که ماهیت سبک وزن MobileNetV3 آن را برای محیط‌های با منابع محدود، مانند برنامه‌های پزشکی در زمان واقعی، مناسب می‌سازد. با استفاده از ماژول‌های توجه، این مدل عملکرد خوبی در دقت طبقه‌بندی در بزرگنمایی‌های مختلف به دست آورد و بهترین دقت آن ۹۶/۲ درصد در بزرگنمایی 40X بود [۱۷].

توگاکار و همکاران یک سازوکار توجه به نام CBAM را در معماری BreastNet که ارایه دادند قرار دادند، بلوک CBAM برای اصلاح ویژگی‌ها به‌طور تطبیقی از طریق پردازش نقشه‌های ویژگی در ابعاد فضایی و کانالی استفاده می‌کند که دقت مدل و کارایی محاسباتی آن را افزایش می‌دهد. علاوه بر CBAM، در BreastNet اجزای مختلفی مانند بلوک‌های پیچشی، چگال، و باقیمانده همراه با تکنیک ابرستون^{۱۷} ادغام شده است. هر یک از این اجزا باعث بهبود تمایز ویژگی‌ها، تقویت جریان گرادیان، بهینه‌سازی انتخاب ویژگی‌ها می‌شوند و در نهایت امکان تحلیل مقیاس‌های مختلف از طریق روش ابرستون را فراهم می‌کند. در مقایسه با مدل‌های پیشرفته مانند AlexNet، VGG-16 و VGG-19، BreastNet عملکرد خیلی خوبی نشان داد و دقت‌هایی بین ۹۵/۸۸ درصد تا ۹۸/۵۲ درصد را در بزرگنمایی‌های مختلف به دست آورد [۱۳].

زو و همکاران با ارائه مدل AHoNet یک چارچوب نوآورانه ارائه دادند که در این مدل سازوکار توجه و نمایش آماری مرتبه بالا را در شبکه‌های هم‌آمیختی باقیمانده ادغام کردند. این مدل با استفاده از یک ماژول توجه کانالی کارآمد و نرمال‌سازی ماتریس توان، ویژگی‌های محلی برجسته و نمایش‌های کلی قوی از تصاویر پاتولوژیک را استخراج می‌کند. AHoNet با ارزیابی روی مجموعه داده‌های BreakHis و BACH، به دقت‌های ۹۹/۲۹ درصد و ۸۵ درصد دست یافت و عملکرد خوبی از خود را در تشخیص سرطان پستان نشان داد [۱۴].

در مدل RANN-BCC محققان از شبکه عصبی ResNet به عنوان روش پشتیبانی‌کننده استفاده کردند تا تشخیص سرطان پستان را تسهیل کند. این مدل می‌تواند بدون نیاز به اطلاعات قبلی در مورد وجود ضایعات سرطانی، تصاویر تمام‌صفحه (WSI)^{۱۸} را به دو دسته کارسینوم مجرای مهاجم و کارسینوم مجرای غیر مهاجم طبقه‌بندی کند. نتایج طبقه‌بندی نشان می‌دهد که مدل RANN-BCC با

19-Squeeze-and-Excitation
20-Global Average Pooling

17-Hyper-column
18-Whole Slide Images

۳. روش پیشنهادی

در این بخش به طور کامل روش شناسی مورد استفاده برای طبقه بندی تصاویر هیستوپاتولوژیک پستان را شرح می دهیم، که نیازمند یک رویکرد قوی برای مدیریت الگوهای پیچیده و تنوع ذاتی در تصاویر پزشکی است که شامل توضیح مراحل پیش پردازش و شبکه عصبی هم آمیختی (CNN) ادغام شده با سازوکارهای توجه است.

۳-۱. پیش پردازش

برای تعمیم پذیری بهتر مدل، تصاویر ورودی را توسط یک سری مراحل پیش پردازش، شامل استانداردسازی و افزایش داده^{۲۱} قرار دادیم که در زیر توضیح داده شده اند

- تغییر اندازه تصویر: اندازه تصاویر ورودی به 256×256 پیکسل تغییر دادیم. این وضوح، تعادلی بین کارایی محاسباتی و حفظ جزئیات کافی برای ثبت الگوهای هیستوپاتولوژیک ایجاد می کند.
- نرمال سازی روشنایی: مقادیر پیکسل را با تقسیم بر ۲۵۵ به بازه $[0, 1]$ نرمال کردیم که این استانداردسازی، همگرایی سریع تر در طول آموزش را تسهیل می کند و توزیع یکنواخت داده ها در کانال ها را تضمین می کند.
- افزایش داده: به منظور افزایش تنوع مجموعه داده، افزایش قدرت تعمیم، و کاهش بیش برآزش، تکنیک های افزایش داده مطابق جدول ۱ مورد استفاده قرار گرفت.
- لایه تغییر مقیاس: یک لایه تغییر مقیاس پس از افزایش داده اعمال شد تا اطمینان حاصل شود که مقادیر ورودی نهایی در محدوده $[0, 1]$ قرار دارند، که این امر را برای حفظ سازگاری داده ها انجام دادیم.

۳-۲. معماری های CNN با سازوکارهای توجه

سه سازوکار توجه به صورت محزا در معماری CNN گنجانده شدند تا تأثیر آن ها بر عملکرد طبقه بندی بررسی

مطالعات دیگری با ارایه مدل های جدیدی از جمله EfficientNetV2s توانستند به درصد تشخیص قابل قبولی (۸۹/۹۱) برسند [۱۸]. مالکی و همکاران در سال ۲۰۲۳ با روش یادگیری انتقالی و بهره گیری از مدل DenseNet201 و سپس ترکیب آن با XGBoost به عنوان طبقه بندی کننده نهایی به دقت تشخیص ۹۱/۹۲ رسیدند [۱۹]. در سال ۲۰۲۴ محققان در پژوهشی با اضافه کردن توجه رو به جلو به معماری VGG16 و ارایه مدلی تحت عنوان FA-VGG16 به نرخ تشخیص ۹۰/۴ در تشخیص دوکلاسه رسیدند [۲۰]. در پژوهشی که سال ۲۰۲۴ توسط Abdulaal و همکاران انجام شد با بهینه سازی معماری GoogleNet و اضافه کردن سازوکار توجه به معماری مربوطه نتایج خوبی در تشخیص سرطان پستان به دست آوردند [۲۱].

مطالعاتی که در بالا بررسی شدند، پیشرفت های قابل توجهی را در استفاده از شبکه های عصبی هم آمیختی و سازوکارهای توجه برای تشخیص هیستوپاتولوژیک سرطان پستان نشان می دهند. این پژوهش ها به چالش های مختلفی مانند استخراج ویژگی های چند مقیاسی، عدم تعادل طبقاتی و ادغام اطلاعات زمینه ای جهانی و محلی پرداختند و به دقت قابل قبولی در طبقه بندی تصاویر دست یافتند. با این حال، محدودیت های خاصی باقی می ماند که می توان به نیاز به چارچوب های محاسباتی کارآمدتر، تعمیم بهتر در میان مجموعه های داده متنوع، و بهبود تفسیرپذیری تصمیمات مدل اشاره کرد. علاوه بر این، در حالی که سازوکارهای توجه مانند PANet، DANet و CBAM در افزایش نمایش ویژگی ها نویدبخش بوده اند، اثربخشی و سازگاری نسبی آنها با تصاویر هیستوپاتولوژیک نیاز به بررسی بیشتر دارد. در این مقاله، ما این شکاف ها را با پیشنهاد یک معماری CNN سفارشی و ارزیابی دقیق عملکرد آن در صورت ادغام با سه سازوکار توجه متمایز: PANet، DANet و CBAM، بررسی می کنیم. نتایج تجربی ما بینش هایی را در مورد نقاط قوت و محدودیت های هر سازوکار توجه در زمینه تشخیص هیستوپاتولوژیک سرطان پستان ارائه می دهد.

جدول ۱: انواع افزایش داده اعمال شده روی تصاویر هیستوپاتولوژی در روش پیشنهادی

توضیحات	تکنیک افزایش داده
وارونگی‌های افقی و عمودی برای شبیه‌سازی تغییرات جهت نمونه‌ها اعمال شدند.	وارونگی تصادفی
تصاویر تا ± 20 درجه چرخیده شدند تا تغییرات احتمالی در چرخش تصویر را پوشش دهند.	چرخش تصادفی
تصاویر تا $\pm 20\%$ درصد بزرگ‌نمایی شدند تا تفاوت‌های سطح بزرگ‌نمایی را جبران کنند.	بزرگ‌نمایی تصادفی
تنظیمات کنتراست برای بهبود یادگیری از نمونه‌هایی با لکه‌گذاری متنوع اعمال شد.	تنظیم تصادفی کنتراست
روشنایی $\pm 20\%$ تنظیم شد، و اشباع در بازه $[0.8, 1.2]$ تغییر داده شد تا تغییرات در نورپردازی و لکه‌گذاری را شبیه‌سازی کند.	تنظیم تصادفی روشنایی و اشباع

که هم اطلاعات معنایی و هم جزئیات فضایی به طور موثر استفاده شوند [۱۱].

۳-۲-۲-۳ DANet (شبکه توجه دوگانه)

DANet از سازوکارهای توجه موقعیتی و کانالی برای ثبت وابستگی‌های مکانی و بین کانالی استفاده می‌کند. بلوک شبکه توجه دوگانه (DANetBlock) برای بهبود بازنمایی ویژگی‌ها با گرفتن وابستگی‌های مکانی و کانالی در نقشه ویژگی ورودی طراحی شده است و این بلوک شامل دو ماژول کلیدی می‌باشد؛ ماژول توجه مکانی^{۲۲} و ماژول توجه کانالی^{۲۳}، این ماژول‌ها به صورت موازی عمل می‌کنند و با هم ترکیب می‌شوند تا یک نقشه ویژگی غنی‌شده تولید کنند که شامل اطلاعات مکانی و کانالی است که ماژول توجه مکانی (PAM) بر مدل‌سازی روابط مکانی در نقشه ویژگی تمرکز دارد، این ماژول ابتدا با استفاده از هم‌آمیختی‌های 1×1 ، سه بازنمایی از نقشه ویژگی ورودی ایجاد می‌کند که عبارتند از پرس‌وجو^{۲۴}، کلید^{۲۵}، و مقدار^{۲۶} که بازنمایی‌های پرس‌وجو و کلید در بعد کانالی کاهش می‌یابند تا کارایی محاسباتی بهبود یابد سپس این بازنمایی‌ها تغییر شکل داده و در هم ضرب می‌شوند تا یک نقشه توجه مکانی محاسبه شود که روابط بین تمام جفت‌های موقعیت‌های مکانی را ثبت می‌کند. یک تابع SoftMax برای نرمال‌سازی وزن‌های توجه اعمال می‌شود تا اطمینان حاصل شود که مجموع آن‌ها برابر با یک است.

22-PAM
23-CAM
24-Query
25-Key
26-Value

شود. هر سازوکار و نحوه ادغام آن در چارچوب CNN در زیر شرح داده شده است:

۳-۲-۱ PANet (شبکه تجمیع مسیر)

PANetBlock یک جزء اصلی از معماری شبکه تجمیع مسیر (PANet) است که برای انجام ادغام دوطرفه ویژگی‌ها بین نقشه‌های ویژگی سطح بالا و سطح پایین طراحی شده است، بلوک PANet را در شکل ۱ مشاهده می‌کنید. این بلوک دو ورودی می‌گیرد: x_1 ، یک نقشه ویژگی سطح بالا که حاوی اطلاعات معنایی غنی اما با وضوح فضایی پایین‌تر است، و x_2 ، یک نقشه ویژگی سطح پایین که جزئیات فضایی دقیق‌تری دارد اما از نظر محتوای معنایی ضعیف‌تر است. این بلوک ابتدا تعداد کانال‌های x_1 و x_2 را با استفاده از هم‌آمیختی‌های 1×1 تنظیم می‌کند تا سازگاری لازم برای ادغام ویژگی‌ها فراهم شود سپس برای ادغام از بالا به پایین، x_1 از بالا به پایین نمونه‌برداری می‌شود تا وضوح آن با x_2 مطابقت پیدا کند، سپس به x_2 اضافه شده و با استفاده از یک هم‌آمیختی 1×1 بهبود می‌یابد. برای ادغام از پایین به بالا، x_2 از پایین به بالا نمونه‌برداری می‌شود تا وضوح آن با x_1 مطابقت پیدا کند، سپس به x_1 اضافه شده و با استفاده از یک هم‌آمیختی 3×3 بهبود می‌یابد. این بلوک دو نقشه ویژگی ادغام شده را خروجی می‌دهد: x_{1_fused} که اطلاعات معنایی سطح بالا را با جزئیات سطح پایین ترکیب می‌کند، و x_{2_fused} که ویژگی‌های سطح بالا را با جزئیات فضایی غنی‌تر می‌کند. این سازوکار ادغام دوطرفه اطمینان حاصل می‌کند

متعادل کند. این رویکرد توجه دوگانه به DANetBlock امکان می‌دهد تا ویژگی‌های غنی و وابسته به زمینه را ثبت کند و منجر به بهبود عملکرد در کارهایی شود که نیاز به درک دقیق ویژگی‌ها دارند [۱۲].

۳-۲-۳. CBAM (ماژول توجه بلوک هم‌آمیختی)

ماژول توجه بلوک هم‌آمیختی^{۲۷} برای بهبود بازنمایی ویژگی‌ها با اعمال متوالی سازوکارهای توجه کانالی و توجه مکانی طراحی شده است و این ماژول شامل دو بخش کلیدی می‌باشد: ماژول توجه کانالی و ماژول توجه مکانی. این بخش‌ها به صورت متوالی عمل می‌کنند تا یک نقشه ویژگی غنی‌شده تولید کنند که شامل هر دو زمینه کانالی و مکانی باشد. ماژول توجه کانالی بر گرفتن وابستگی‌های کانالی تمرکز دارد که در ابتدا با استفاده از هر دو گزینه میانگین‌گیری سراسری و بیشینه‌گیری سراسری، اطلاعات مکانی را از نقشه ویژگی ورودی جمع‌آوری می‌کند. سپس این توصیف‌گرهای جمع‌آوری شده، اطلاعات مکملی درباره زمینه سراسری نقشه ویژگی ارائه می‌دهند و خروجی‌های جمع‌آوری شده سپس به یک شبکه چندلایه‌ای مشترک^{۲۸} با یک لایه پنهان وارد می‌شوند. شبکه چندلایه‌ای مشترک بعد کانالی را با ضریب کاهش^{۲۹} (مثلاً ۸) کاهش می‌دهد تا پیچیدگی محاسباتی کمتر شود و سپس آن را به بعد اصلی باز می‌گرداند. خروجی‌های با شبکه چندلایه‌ای مشترک هم ترکیب شده و از طریق یک فعال‌سازی سیگموئید عبور می‌کنند تا وزن‌های توجه کانالی تولید شوند که بین ۰ و ۱ نرمال‌سازی شده‌اند سپس این وزن‌ها به نقشه ویژگی ورودی اعمال می‌شوند تا کانال‌های مهم تأکید شوند و کانال‌های کم‌اهمیت سرکوب شوند، که نتیجه آن یک نقشه ویژگی بهبودیافته کانالی است.

ماژول توجه مکانی بر گرفتن وابستگی‌های مکانی تمرکز می‌کند این ماژول با گرفتن خروجی ماژول توجه کانالی اطلاعات کانالی را با استفاده از میانگین‌گیری کانالی و

سپس وزن‌های توجه برای تجمیع بازنمایی مقدار استفاده می‌شوند و یک نقشه ویژگی بهبودیافته مکانی تولید می‌کنند. در نهایت، خروجی با ورودی اصلی از طریق یک اتصال باقیمانده ترکیب می‌شود و با یک پارامتر یادگیری شده به نام گاما مقیاس می‌شود تا ویژگی‌های اصلی حفظ شوند و در عین حال زمینه مکانی نیز اضافه شود.

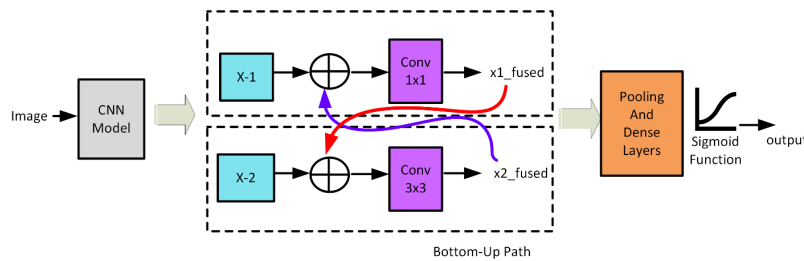
ماژول توجه کانالی بر مدل‌سازی روابط کانالی تمرکز دارد. این ماژول نقشه ویژگی ورودی را تغییر شکل می‌دهد تا یک نقشه توجه کانالی محاسبه کند که روابط بین تمام جفت‌های کانال‌ها را ثبت می‌کند به این شکل که ورودی ابتدا تغییر شکل داده و جابه‌جا می‌شود تا ابعاد کانالی تراز شوند، ضرب ماتریسی برای محاسبه شباهت بین کانال‌ها استفاده می‌شود، و یک تابع SoftMax برای نرمال‌سازی وزن‌های توجه اعمال می‌شود سپس این وزن‌ها برای تجمیع نقشه ویژگی ورودی در بعد کانالی استفاده می‌شوند و یک نقشه ویژگی بهبودیافته کانالی تولید می‌کنند. مشابه PAM، خروجی با ورودی اصلی از طریق یک اتصال باقیمانده ترکیب می‌شود و با یک پارامتر یادگیری شده به نام گاما مقیاس می‌شود تا ویژگی‌های اصلی حفظ شوند و در عین حال زمینه کانالی نیز اضافه شود.

خروجی‌های PAM و CAM با هم جمع می‌شوند تا خروجی نهایی DANetBlock تولید شود و در شکل ۲ قابل ملاحظه است. این نقشه ویژگی ترکیبی شامل اطلاعات بهبودیافته مکانی و کانالی است که به شبکه اجازه می‌دهد تا وابستگی‌های پیچیده در داده‌ها را ثبت کند. با ادغام این دو نوع سازوکار توجه، DANetBlock به طور موثر قدرت بازنمایی نقشه ویژگی را بهبود می‌بخشد و آن را برای کارهای بعدی مانند تقسیم‌بندی یا طبقه‌بندی مناسب‌تر می‌کند. استفاده از اتصالات باقیمانده نیز این اطمینان را می‌دهد که ویژگی‌های اصلی حفظ می‌شوند، در حالی که پارامترهای مقیاس یادگیری شده (گاما) به شبکه اجازه می‌دهند تا سهم سازوکارهای توجه را به صورت تطبیقی

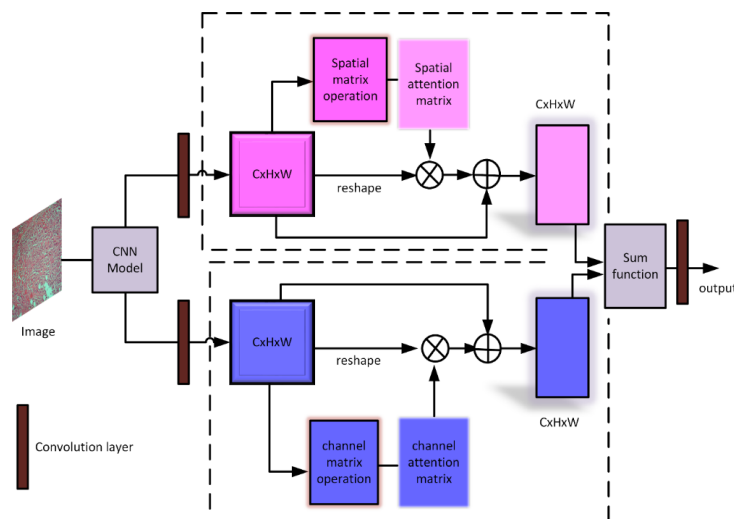
27-CBAM

28-MLP

29-Reduction ratio



شکل ۱. طرح بلوک PANet



شکل ۲. طرح کلی بلوک DANet

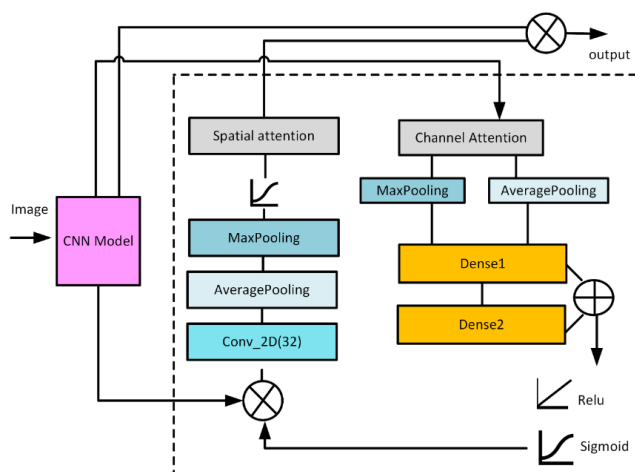
نقشه ویژگی غنی شده است که شامل هر دو زمینه کانالی و مکانی است و آن را برای کارهای بعدی مانند طبقه‌بندی، تقسیم‌بندی یا تشخیص اشیاء مناسب‌تر می‌کند. استفاده از میانگین‌گیری سراسری و شبکه‌های چندلایه‌ای مشترک، کارایی محاسباتی را تضمین می‌کند، در حالی که اعمال متوالی سازوکارهای توجه اطمینان می‌دهد که هر دو نوع زمینه به طور مؤثر در نظر گرفته شوند. این رویکرد توجه دوگانه به مازول توجه بلوک هم‌آمیختگی امکان می‌دهد تا ویژگی‌های غنی و وابسته به زمینه را ثبت کند و منجر به بهبود عملکرد در کارهایی شود که نیاز به درک دقیق ویژگی‌ها دارند [۱۳].

۴. آزمایش‌ها و نتایج

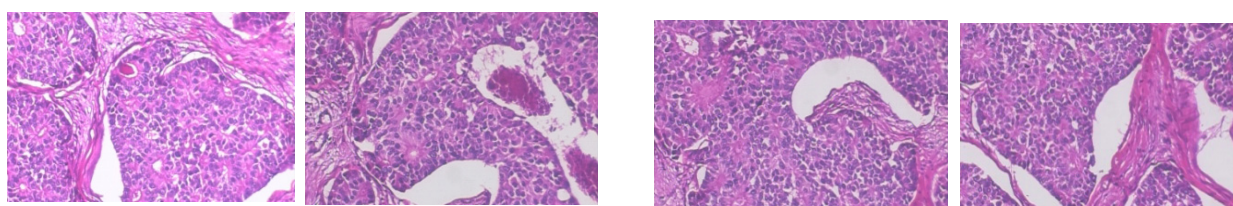
در این بخش پس از معرفی مجموعه داده و معیارهای کارایی استاندارد، شبکه‌های مبتنی بر سازوکار توجه معرفی شده با روش‌های موجود مقایسه و ارزیابی

بیشینه‌گیری کانالی جمع‌آوری می‌کند. این توصیف‌گرهای جمع‌آوری‌شده، اطلاعات مکملی درباره زمینه مکانی نقشه ویژگی ارائه می‌دهند سپس دو توصیف‌گر در بعد کانالی به هم متصل می‌شوند تا یک نقشه ویژگی دوکانالی تشکیل دهند. سپس یک لایه هم‌آمیختگی 7×7 با فعال‌سازی سیگموئید اعمال می‌شود تا وزن‌های توجه مکانی محاسبه شوند و بین ۰ و ۱ نرمال‌سازی شوند در نهایت این وزن‌ها به نقشه ویژگی اعمال می‌شوند تا مکان‌های مهم تأکید شوند و مکان‌های کم‌اهمیت سرکوب شوند، که نتیجه آن یک نقشه ویژگی بهبودیافته مکانی است.

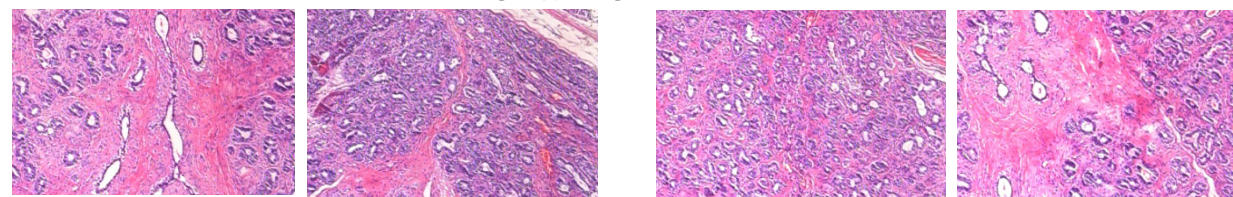
همان‌طور که در شکل ۳ مشخص است خروجی‌های مازول توجه کانالی و مازول توجه مکانی به صورت متوالی به نقشه ویژگی ورودی اعمال می‌شوند و مازول توجه کانالی نقشه ویژگی را با تأکید بر کانال‌های مهم بهبود می‌بخشد، در حالی که مازول توجه مکانی آن را با تأکید بر مکان‌های مهم بیشتر بهبود می‌بخشد. خروجی نهایی یک



شکل ۳. طرح کلی بلوک CBAM



الف) تصاویر سرطان بدخیم پستان



ب) تصاویر سرطان خوش خیم پستان

شکل ۴. تصاویر ردیف بالا مربوط به سرطان بدخیم پستان و تصاویر ردیف پایین مربوط به سرطان خوش خیم پستان

می‌شوند.

از اسلایدهای نمونه‌برداری شده از بافت سینه در چهار سطح بزرگنمایی مختلف ($40\times$ ، $100\times$ ، $200\times$ و $400\times$) تهیه شده‌اند. فرآیند آماده‌سازی نمونه‌ها شامل رنگ‌آمیزی با هماتوکسیلین و ائوزین و استفاده از فرایند استاندارد پارافین بوده است. تصاویر با استفاده از یک دستگاه سامسونگ با تفکیک‌پذیری بالا و سیستم میکروسکوپی Olympus BX-50 با لنز $33\times$ ثبت شده‌اند. این تصاویر که دارای رنگ‌بندی قرمز-سبز-آبی هستند، در ابعاد 700×600 پیکسل ضبط شده‌اند. مجموعه داده BreakHis اطلاعات جامعی از تصاویر خوش‌خیم و بدخیم ارائه می‌دهد و برای تحلیل‌های یادگیری عمیق در تشخیص سرطان مناسب است [۲۲].

۴-۱. مجموعه داده استاندارد

برای ارزیابی عملکرد مدل توسعه‌یافته، از مجموعه داده BreakHis که به‌طور عمومی در دسترس است، استفاده کردیم. این مجموعه شامل ۷۹۰۹ تصویر پاتولوژی از بافت سرطان سینه است که از ۸۲ بیمار ناشناس جمع‌آوری شده و پراکندگی آن در جدول ۲ آورده شده است. تصاویر موجود در این مجموعه به دو دسته اصلی تقسیم می‌شوند که در آن ۲۴۸۰ تصویر مربوط به تومورهای خوش‌خیم و ۵۴۲۹ تصویر مربوط به تومورهای بدخیم است و نمونه‌هایی از هر دسته در شکل ۴ قابل مشاهده است. این تصاویر

(FP) نتایج مهم یا حساسی داشته باشد، کاربرد ویژه‌ای دارد. این معیار تضمین می‌کند که پیش‌بینی‌های مدل برای کلاس مثبت، دقیق و قابل اطمینان هستند، که به‌ویژه در کاربردهایی مانند تشخیص پزشکی، اهمیت بالایی دارد.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

● **فراخوانی (Recall):** فراخوانی، که به‌عنوان نرخ صحیح مثبت (TPR) نیز شناخته می‌شود، درصد نمونه‌های مثبت واقعی شناسایی شده توسط مدل را اندازه‌گیری می‌کند. این معیار برای ارزیابی مدل‌هایی که کاهش منفی‌های کاذب (FN) اهمیت دارد، به‌طور خاص مفید است. این معیار در کاربردهایی که شناسایی موارد مثبت حیاتی است، مانند تشخیص بیماری، بسیار پرکاربرد است حتی اگر منجر به افزایش مثبت‌های کاذب شود، شناسایی حداکثری مثبت‌های واقعی را تضمین می‌کند [۲۴].

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

● **امتیاز (F1-Score):** امتیاز F1، که ترکیبی از دقت مثبت و فراخوانی است، میانگین هارمونیک این دو معیار را ارائه می‌دهد. این معیار برای شرایطی که نیاز به تعادل میان دقت مثبت و فراخوانی وجود دارد، مناسب است.

$$F1 - Score = \frac{Precision \cdot Recall}{Precision + Recall} \quad (4)$$

● **AUC-ROC:** مساحت زیر منحنی مشخصه عملکرد گیرنده (AUC-ROC) توانایی مدل در تمایز بین کلاس‌های مثبت و منفی را در سطوح آستانه مختلف ارزیابی می‌کند. مقدار AUC برابر با ۱ نشان‌دهنده عملکرد بی‌نقص است، در حالی که مقدار ۰/۵ به معنی عدم توانایی تفکیک است. مقادیر بالاتر AUC به عملکرد بهتر مدل اشاره دارد و این معیار به‌ویژه برای مقایسه مدل‌های مختلف بسیار مفید است [۲۴].

$$AUC - ROC = \int_0^1 (TPR)d(FPR) \quad (5)$$

که در آن TPR و FPR به ترتیب نرخ مثبت واقعی و نرخ

جدول ۲. پراکندگی دیتاست Breakhis

مجموع	بدخیم	خوش خیم	بزرگنمایی
۱۹۹۵	۱۳۷۰	۶۲۵	۴۰X
۲۰۸۱	۱۴۳۷	۶۴۴	۱۰۰X
۲۰۱۳	۱۳۹۰	۶۲۳	۲۰۰X
۱۸۲۰	۱۲۳۲	۵۸۸	۴۰۰X
۷۹۰۹	۵۴۲۹	۲۴۸۰	مجموع
۸۲	۵۸	۲۴	تعداد بیماران

۴-۲. معیارهای ارزیابی کارایی

برای ارزیابی و مقایسه کارایی روش‌های مختلف یادگیری ماشین، ضروری است که از معیارهای استاندارد استفاده کنیم. این معیارها به ما کمک کردند تا عملکرد و کیفیت پیش‌بینی‌های مدل‌ها را به‌طور دقیق بررسی کنیم. در ادامه مهم‌ترین معیارهای ارزیابی کارایی روش‌های دسته‌بندی کننده معرفی می‌شوند.

● **دقت (Accuracy):** دقت معیاری پایه‌ای برای بررسی عملکرد مدل‌های طبقه‌بندی است و نسبت پیش‌بینی‌های درست (چه برای کلاس مثبت و چه منفی) به کل پیش‌بینی‌ها را نشان می‌دهد. این معیار به‌طور کلی بیان می‌کند که مدل در پیش‌بینی کلی خود چند درصد درست عمل کرده است [۲۳].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

که در آن TP, TN, FP, FN به ترتیب تعداد پیش‌بینی‌های منفی کاذب، مثبت کاذب، منفی واقعی و مثبت واقعی هستند. این معیار به دلیل سادگی و وضوحش، برای ارزیابی مدل‌های با داده‌های متوازن بین کلاس‌ها بسیار مناسب است. اما در مواجهه با داده‌های نامتوازن، ممکن است نتایج گمراه‌کننده ارائه دهد، زیرا ممکن است مدلی با تمرکز بر پیش‌بینی کلاس اکثریت، دقت ظاهراً بالایی داشته باشد.

● **دقت مثبت (Precision):** دقت مثبت، درصد نمونه‌های پیش‌بینی شده مثبت که واقعاً مثبت بوده‌اند را بیان می‌کند. این معیار به‌طور خاص در شرایطی که وقوع مثبت‌های کاذب

مثبت کاذب هستند.

● **ماتریس درهم‌ریختگی (Confusion Matrix):** ماتریس درهم‌ریختگی یک نمایش جدولی است که برای ارزیابی عملکرد مدل‌های طبقه‌بندی استفاده می‌شود. این ماتریس یک تجزیه و تحلیل دقیق از پیش‌بینی‌های مدل در مقایسه با برچسب‌های واقعی را ارائه می‌دهد و برای درک انواع خطاهای مدل بسیار مفید است. برای یک مسئله طبقه‌بندی دودویی، ماتریس درهم‌ریختگی از چهار جزء کلیدی تشکیل شده است TP, TN, FP, FN که به ترتیب تعداد پیش‌بینی‌های منفی کاذب، مثبت کاذب، منفی واقعی و مثبت واقعی هستند. از ماتریس درهم‌ریختگی، معیارهای ارزیابی مختلفی مانند دقت (Accuracy)، صحت (Precision)، فراخوانی (Recall) و F1 امتیاز قابل محاسبه است.

● **مساحت زیر منحنی دقت-فراخوانی (AUC-PR):** مساحت زیر منحنی دقت-فراخوانی (AUC-PR) یک معیار عملکردی است که برای ارزیابی کیفیت مدل‌های طبقه‌بندی، به‌ویژه در مواردی که داده‌ها نامتوازن هستند، استفاده می‌شود. برخلاف ROC-AUC که نرخ مثبت‌های واقعی (فراخوانی) را در برابر نرخ مثبت‌های کاذب رسم می‌کند، منحنی دقت-فراخوانی بر رابطه بین صحت (Precision) و فراخوانی (Recall) در آستانه‌های طبقه‌بندی مختلف تمرکز دارد. هرچه AUC-PR بالاتر باشد، عملکرد مدل بهتر است، به‌طوری که مقدار ۱ نشان‌دهنده صحت و فراخوانی کامل و مقدار ۰.۵ نشان‌دهنده یک مدل تصادفی است.

داده‌های آموزش به عنوان داده اعتبارسنجی استفاده شد. این تقسیم‌بندی تصادفی به مدل کمک می‌کند تا تعمیم‌پذیری خوبی برای داده‌های دیده نشده داشته باشد و از بروز سوگیری‌هایی که ممکن است از تقسیمات غیرتصادفی ناشی شود، جلوگیری کند.

برای بهینه‌سازی عملکرد مدل، فرآیند تنظیم سیستماتیک ابرپارامترها انجام گرفت و ابرپارامترهای کلیدی مانند نرخ یادگیری ۰.۰۰۱، اندازه دسته ۳۲، تعداد دوره ۳۳ و نرخ افت ۰.۰۰۱ به دقت تنظیم شدند. روش Hyperband برای جستجوی مؤثر و انتخاب بهترین ترکیب ابرپارامترها مورد استفاده قرار گرفت [۲۵]. در نتیجه این تنظیمات، نرخ یادگیری بهینه برابر ۰.۰۰۰۴ و اندازه دسته برابر ۱۶ تعیین شد، که بهترین دقت اعتبارسنجی را ارائه دادند. همچنین، تعداد دوره‌ها بین ۱۰ تا ۵۰ متغیر بود تا مدت زمان بهینه آموزش مدل مشخص شود. پس از ارزیابی عملکرد مدل در این بازه، تعداد ۳۰ دوره به‌عنوان نقطه تعادل بین همگرایی و جلوگیری از بیش‌برازش انتخاب شد. در این مرحله، کاهش اعتبارسنجی به حالت پایدار رسید، که نشان‌دهنده یادگیری مؤثر مدل بدون افت عملکرد قابل‌توجه بر روی داده‌های دیده نشده است. در این مدل، نرخ افت ۰.۰۰۱ انتخاب شد که تعادل مناسبی بین جلوگیری از بیش‌برازش و حفظ ظرفیت یادگیری فراهم کرد. این روش‌ها و تنظیمات پارامترها به‌گونه‌ای انتخاب شدند که عملکرد مدل، زمان آموزش و کارایی محاسباتی بهینه شوند.

۴-۴. ارزیابی و مقایسه کارایی

در جدول ۳ نتایج مدل‌های مختلف بر روی مجموعه داده استاندارد BreakHis با نتایج به دست آمده از سه مدل DANet, PANet و CBAM با استفاده از پنج معیار کلیدی دقت، صحت، فراخوانی، امتیاز F1 و مساحت زیر منحنی ROC مقایسه شده است. برای مقایسه بهتر، برترین

۴-۳. تنظیمات آزمایشی

در این مطالعه، چند مدل یادگیری عمیق مبتنی بر سازوکار توجه، برای تشخیص سرطان سینه پیاده‌سازی شد. مجموعه داده استاندارد BreakHis به صورت تصادفی به دو بخش آموزش و آزمایش تقسیم شد، به‌گونه‌ای که ۸۰ درصد داده‌ها برای آموزش انتخاب و ۲۰ درصد باقی‌مانده تکرار شد که میانه نتایج گزارش شد و ۲۰ درصد باقی‌مانده برای آزمایش در نظر گرفته شدند. همچنین از ۲۰ درصد

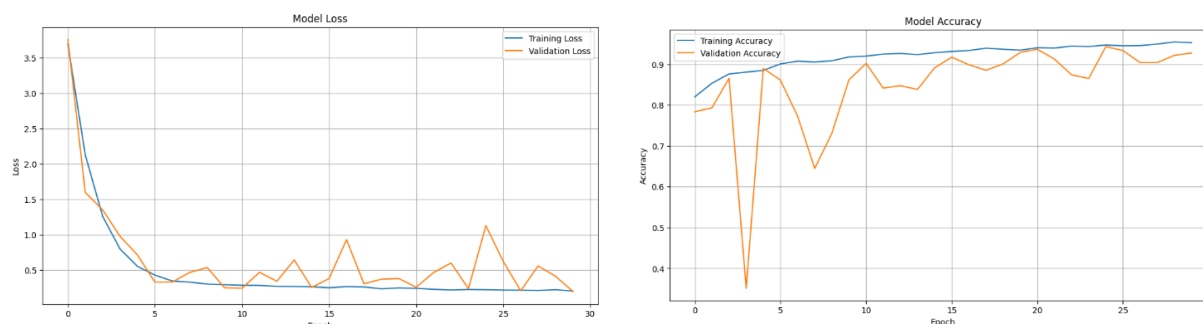
جدول ۳. نتایج مدل‌های مختلف براساس معیارهای کارایی مختلف روی مجموعه داده Breakhis

مدل‌ها	Accuracy (%)	Precision (%)	Recall (%)	F1_score (%)	ROC-AUC
MobileNetV3	۸۹/۵۲	۸۸	۸۸/۵۱	۸۷/۸	---
Resnet18	۸۳/۱۳	۸۴/۹۲	۸۰/۲۵	۸۲/۱۲	---
EfficientNetB0	۸۷/۴۲	۸۶/۲۳	۸۵/۲۳	۸۵/۳۹	---
MobileNetV3_SCA	۹۴/۸	۹۳/۸۰	۹۴/۹۱	۹۴/۸۳	---
EfficientNetV2-S	۸۹/۹۱	۹۰/۲۳	۸۹/۸۵	۷۹/۵۵	---
DenseNet201-XGboost	۹۱/۹۲	---	---	---	---
FA-VGG16	۹۰/۴	---	---	---	---
GoogleNet	۹۴/۲۲	۹۲/۵۳	۹۳/۶	۹۳/۱۲	---
RANN-BCC	۹۲/۴۵	۹۱/۰۰	۹۸/۰۰	۹۴/۰۰	---
CBAM	۹۲/۷۶	۹۴/۲۱	۹۵/۳۶	۹۴/۷۸	۹۷/۱۷
PANet	۷۷/۳۸	۹۳/۳۸	۷۳/۰۸	۸۱/۹۹	۸۷/۱۰
DANet	۹۳/۰۲	۹۵/۴۳	۹۴/۶۳	۹۵/۰۳	۹۷/۷۰

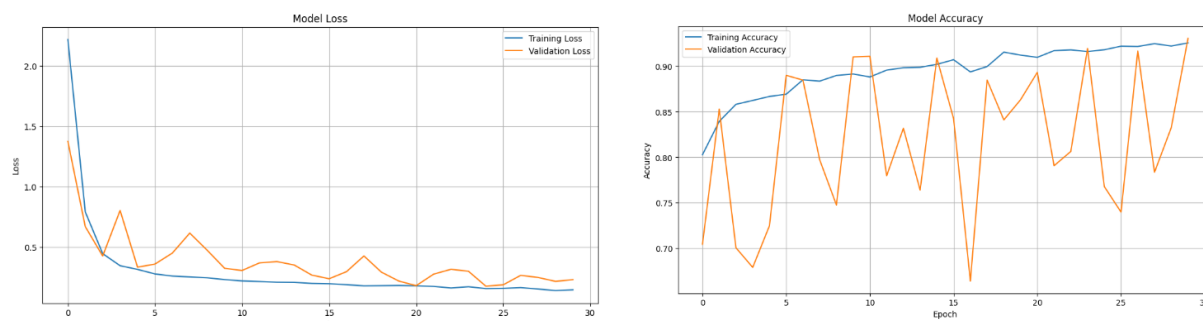
۴-۵. مقایسه روند آموزش

نمودار دقت و خطا در روند آموزش سه مدل MobileNetV3، Resnet18، EfficientNetB0، MobileNetV3_SCA، EfficientNetV2-S، DenseNet201-XGboost، FA-VGG16، GoogleNet، RANN-BCC، CBAM، PANet و DANet در شکل ۵ روند مشابهی را برای سه سازوکار توجه مختلف نشان می‌دهند. در مراحل اولیه آموزش، دقت هر مدل نسبتاً پایین بود، چرا که مدل در حال یادگیری الگوهای پایه‌ای موجود در داده‌ها بود. با گذشت زمان و طی ۳۰ دوره، دقت به تدریج افزایش یافت، که نشان‌دهنده بهبود پیوسته مدل در طبقه‌بندی صحیح داده‌ها بود. در فرآیند آموزش از توقف زودهنگام^{۳۵} استفاده شد، که آموزش را در دوره ۳۰ام متوقف کرد. این توقف زمانی رخ داد که دقت مدل به یک سطح ثابت رسید و تغییرات بیشتری در بهبود دقت مشاهده نشد. توقف زودهنگام از بیش‌برازش جلوگیری کرده و زمان آموزش را بهینه کرد، زیرا فرآیند زمانی متوقف شد که عملکرد روی داده‌های اعتبارسنجی دیگر بهبود نمی‌یافت. مدل CBAM به طور پیوسته در حال افزایش درستی و همچنین کاهش خطا بود در مدل DANet نیز خطا در حال کمتر شدن اما درستی به صورت نوسانی در حال افزایش بود که در دوره ۳۰ام بیش‌ترین حد خود را تجربه کرد و مدل PANet خطا در حال کاهش و درستی به طور

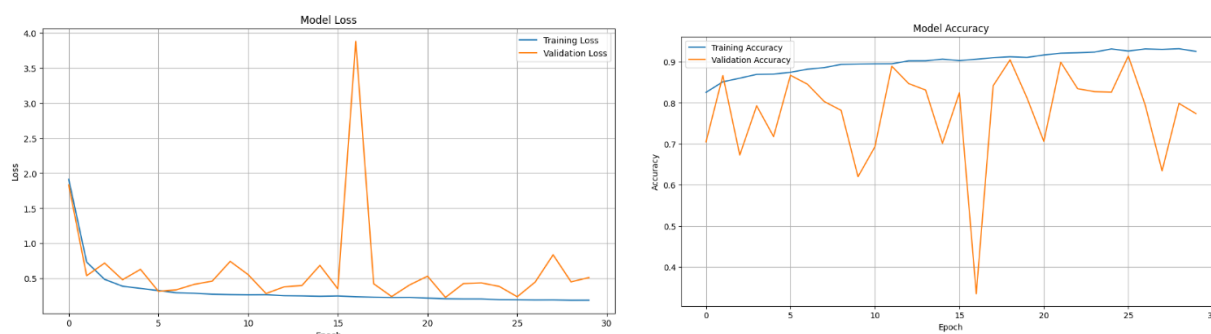
نتیجه بر حسب هر معیار کارایی در هر ستون، به صورت پررنگ نمایش داده شده است. مدل‌های DANet، PANet و CBAM که مبتنی بر سازوکار توجه هستند، به طور کلی در مقایسه با دیگر مدل‌ها برتری‌های قابل توجهی را نشان می‌دهند. با این که از لحاظ دقت کلی، مدل MobileNetV3_SCA با دقت ۹۴/۸ درصد بالاترین عملکرد را دارد و از نظر معیارهای مختلف مانند دقت و یادآوری نیز پیش‌تاز است، اما مدل‌های مبتنی بر سازوکار توجه به دلیل توانایی آن‌ها در تمرکز بر ویژگی‌های مهم و نادیده گرفتن اطلاعات بی‌اهمیت، همچنان نتایج قابل توجهی را در مقایسه با دیگر روش‌ها ارائه می‌دهند. مدل CBAM با دقت ۹۲/۷۶ درصد و فراخوانی ۹۵/۳۶ درصد عملکرد بسیار خوبی دارد و نشان می‌دهد که توانایی تمرکز بر ویژگی‌های کلیدی تصاویر به آن کمک کرده است تا این مدل در یافتن تمامی نمونه‌های مثبت توانا شود. همچنین، مدل DANet با دقت ۹۳/۰۲ درصد و دقت ۹۵/۴۳ درصد توانسته است به خوبی از سازوکار توجه برای تقویت ویژگی‌ها استفاده کند. DANet با عملکرد برتر در ROC-AUC، که نشان از توانایی بهتر این مدل در تفکیک بین دو کلاس، با در نظر گرفتن تعادل میان مثبت واقعی و مثبت کاذب است نیز برتری خود را به اثبات رسانده است.



(الف) خطا و دقت CBAM



(ب) خطا و دقت DANet



(ج) خطا و دقت PANet

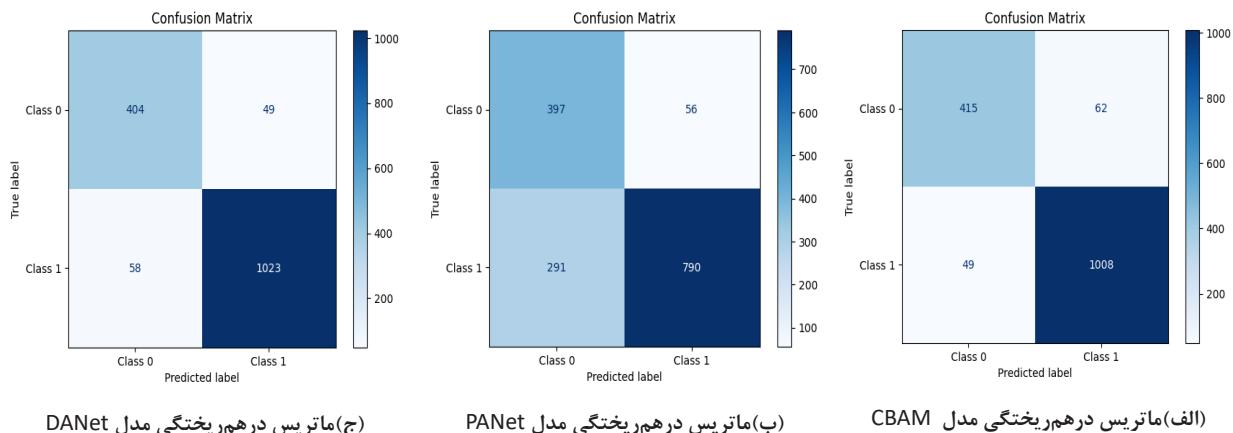
شکل ۵. نمودار خطا و دقت هر سه مدل در روند آموزش

سلول‌های غیرقطری (مثبت کاذب (FP) و منفی کاذب (FN)) پیش‌بینی‌های نادرست را نمایش می‌دهند. در نتایج به‌دست آمده، مدل‌هایی که از سه سازوکار مختلف توجه در شبکه عصبی هم‌آمیختی استفاده کردند، نتایج متفاوتی نشان دادند. در مدل CBAM، تعداد ۴۱۵ منفی واقعی پیش‌بینی شد که نشان‌دهنده تعداد قابل‌توجهی از شناسایی‌های درست برای کلاس منفی است. همچنین، ۱۰۰۸ مثبت واقعی ثبت شد که مؤید توانایی مدل در شناسایی صحیح کلاس مثبت است. با این حال، این مدل ۶۲ مثبت کاذب و ۴۹ منفی کاذب نیز تولید کرد که به ترتیب نشان‌دهنده پیش‌بینی اشتباه

پیوسته در حال افزایش بود بجز در دوره ۱۵ که خطا ناگهانی زیاد و درستی کم شد.

۴-۶. مقایسه ماتریس درهم‌ریختگی

شکل ۶ ماتریس درهم‌ریختگی^{۳۶} مدل پیشنهادی را نشان می‌دهد. هر خانه در این ماتریس تعداد پیش‌بینی‌های مدل بر روی داده‌های آزمایشی را مشخص می‌کند. سلول‌های قطری ماتریس (مثبت واقعی (TP) و منفی واقعی (TN)) نشان‌دهنده پیش‌بینی‌های صحیح هستند، در حالی که



(ج) ماتریس درهم‌ریختگی مدل DANet

(ب) ماتریس درهم‌ریختگی مدل PANet

(الف) ماتریس درهم‌ریختگی مدل CBAM

شکل ۶. ماتریس‌های درهم‌ریختگی سه سازوکار توجه

نمونه‌های مثبت است. بنابراین، DANet تعادل بهتری بین شناسایی‌های صحیح و خطاها نشان می‌دهد و به عنوان بهترین مدل در بین این سه مدل شناخته می‌شود.

۴-۷. منحنی دقت - فراخوانی (PR)

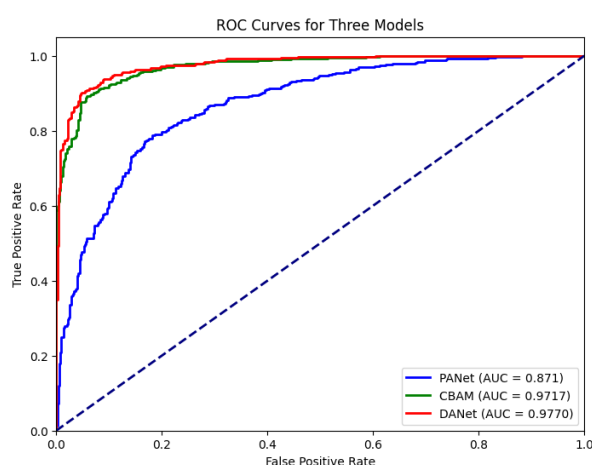
عملکرد مدل شبکه عصبی هم‌آمیختی (CNN) با تمرکز بر موازنه بین دقت و فراخوانی و با استفاده از منحنی‌های دقت فراخوانی (PR) ارزیابی شد. این موضوع به ویژه در مواردی که ممکن است عدم تعادل طبقاتی در مجموعه داده وجود داشته باشد، اهمیت دارد. مساحت زیر منحنی (AUC-PR) برای هر سازوکار توجه محاسبه شد و در شکل ۷ قابل مشاهده است. نتایج ما نشان‌دهنده تفاوت‌های قابل توجهی در عملکرد بین سه سازوکار توجه آزمایش شده است. مدل DANet به AUC-PR قابل توجه ۰/۹۹۰۳ دست می‌یابد، که عملکرد عالی در ایجاد تعادل موثر بین دقت و فراخوانی در آستانه‌های مختلف را نشان می‌دهد. این نتیجه برتر نشان می‌دهد که مدل‌سازی صریح هر دو وابستگی مکانی و کانالی برای این کار خاص بسیار سودمند است. توانایی DANet در یادگیری همزمان این که به کجا توجه کند و به چه چیزی توجه کند، آن را قادر می‌سازد تا مرتبط‌ترین ویژگی‌ها را از تصاویر ورودی شناسایی و استفاده کند. مدل CBAM نیز با قدرت به AUC-PR برابر با ۰/۹۸۶۲ دست می‌یابد. این نتیجه همچنین نشان‌دهنده مفید

نمونه‌های منفی به‌عنوان مثبت و عدم شناسایی برخی نمونه‌های مثبت است.

در مدل DANet، تعداد ۴۰۴ منفی واقعی پیش‌بینی شد که به خوبی شناسایی نمونه‌های منفی را نشان می‌دهد. علاوه بر این، تعداد ۱۰۲۳ مثبت واقعی ثبت شد که نشان‌دهنده دقت بالای مدل در شناسایی مثبت‌ها است. اما این مدل ۴۹ مثبت کاذب و ۵۸ منفی کاذب نیز داشت که نشان می‌دهد در برخی موارد نمونه‌های منفی به‌عنوان مثبت شناسایی شده و یا برخی نمونه‌های مثبت نادیده گرفته شده‌اند. در نهایت، در مدل PANet، تعداد ۳۹۷ منفی واقعی و ۷۹۰ مثبت واقعی پیش‌بینی شد که به ترتیب نشان‌دهنده دقت مناسب در شناسایی نمونه‌های منفی و مثبت است. با این حال، مدل ۵۶ مثبت کاذب و ۲۹۱ منفی کاذب تولید کرد که نشان‌دهنده تعداد قابل توجهی پیش‌بینی اشتباه است، به‌ویژه در شناسایی نمونه‌های مثبت.

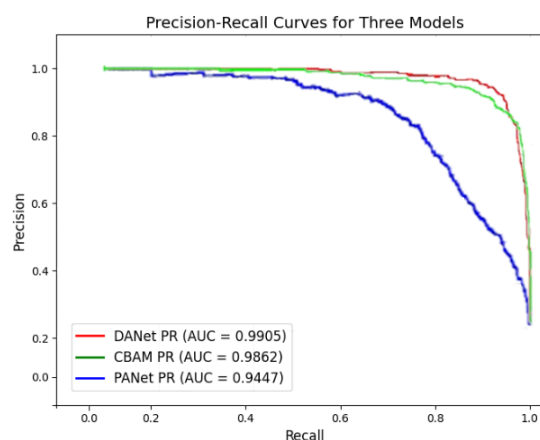
بر اساس نتایج ماتریس درهم‌ریختگی، مدل DANet در مقایسه با سایر مدل‌ها عملکرد بهتری دارد. DANet تعداد بیشتری از مثبت‌های واقعی (۱۰۲۳) و تعداد مناسبی از منفی‌های واقعی (۴۰۴) را به دست آورد که نشان‌دهنده دقت بالاتر در شناسایی صحیح هر دو کلاس مثبت و منفی است. مدل CBAM نیز عملکرد خوبی داشت، اما مثبت‌های کاذب (۶۲) و منفی‌های کاذب (۴۹) آن کمی بیشتر بود. از طرف دیگر، PANet تعداد قابل توجهی منفی کاذب (۲۹۱) داشت که نشان‌دهنده عملکرد ضعیف‌تر آن در شناسایی

گیرنده (ROC) دیدگاه‌های مکملی از عملکرد مدل ارائه می‌دهند. نمودارهای ROC به دست آمده برای هریک از سه مدل بررسی شده، در شکل ۸ نمایش داده شده است. چنان که در قسمت قبل مشاهده شده، مدل DANet بالاترین AUC-PR را با مقدار ۰/۹۹۰۳ به دست آورد، و پس از آن CBAM با ۰/۹۸۶۲ و سپس PANet با ۰/۹۴۴۷ قرار گرفتند. این نشان می‌دهد که سازوکار توجه دوگانه DANet، که به طور صریح وابستگی‌های مکانی و کانالی را مدل‌سازی می‌کند، برای این کار موثرترین است و امکان استفاده بهتر از ویژگی‌ها را فراهم می‌کند. مطابق شکل ۸، تحلیل ROC تا حد زیادی این یافته‌ها را تصدیق می‌کند، به طوری که DANet به AUC-ROC برابر با ۰/۹۷۷۰، CBAM به ۰/۹۷۱۷ و PANet به ۰/۸۷۱۰ دست یافتند. این نتایج نشان می‌دهد که DANet نه تنها بهترین تعادل را بین دقت و فراخوانی ایجاد می‌کند، بلکه در تفکیک بین کلاس‌ها نیز برتری دارد. در کل، روندهای همسان در منحنی‌های PR و ROC برتری DANet را تأیید می‌کند، در حالی که مزایای سازوکارهای توجه را برای بهبود عملکرد مدل برجسته می‌سازد. عملکرد پایین‌تر PANet در هر دو منحنی نشان می‌دهد که رویکرد سلسله مراتبی و چند سطحی ممکن است برای این کار بهینه نباشد.



شکل ۸. نمودار زیر منحنی عملکرد گیرنده (ROC) برای مدل‌های PANet , CBAM , DANet

بودن توجه کانالی و مکانی با انجام متوالی آنها در شبکه است. اگرچه کمی کمتر از DANet است، اختلاف عملکرد نشان می‌دهد که توجه دوگانه ممکن است بر توجه‌های متوالی برای این کار خاص برتری داشته باشد. در نهایت، مدل PANet دارای AUC-PR برابر با ۰/۹۴۴۷ است. این نمره در مقایسه با DANet و CBAM به طور قابل‌توجهی کمتر است بر این اساس، ماهیت سلسله مراتبی توجه هرمی و تمرکز آن بر ویژگی‌های چند سطحی آن به اندازه کافی برای تشخیص در این کاربرد خاص مناسب نیست. با این حال، هر سه رویکرد از CNN به تنهایی بهتر عمل می‌کنند، و این بیانگر این موضوع است که گنجاندن سازوکارهای توجه برای افزایش توانایی مدل در یادگیری بازنمایی‌های معنادار بسیار حائز اهمیت است. نتایج به وضوح برتری DANet را برای این کار خاص نشان می‌دهد و حاکی از آن است که روش توجه دوگانه آن بسیار موثر است. توجه به این نکته مهم است که عملکرد سازوکارهای توجه به شدت به کاربرد وابسته است و تجزیه و تحلیل ویژگی‌های داده‌ها و ویژگی‌های خاص برای ارائه یک توضیح محکم برای اعداد حاصل بسیار مهم است.



شکل ۷. نمودار زیر منحنی دقت-فراخوانی (PR) برای مدل‌های PANet , DANet , CBAM

۴-۸. منحنی مشخصه عملکرد گیرنده (ROC)

منحنی‌های دقت-فراخوانی (PR) و منحنی عملکرد

۵- بحث و نتیجه گیری

در این مطالعه، ما به دنبال بهبود تشخیص هیستوپاتولوژیک سرطان پستان بودیم و سه سازوکار توجه DANet، PANet و CBAM را با یک مدل CNN سفارشی ترکیب کردیم. ارزیابی‌های سیستماتیک نشان داد که DANet با دقت ۹۳/۰۲ درصد و AUC-ROC برابر با ۹۷/۷۰ درصد بهترین عملکرد را در مقایسه با CBAM و PANet داشته است. این برتری به توانایی DANet در مدل‌سازی هم‌زمان وابستگی‌های مکانی و کانالی مربوط می‌شود که به آن اجازه می‌دهد بر مناطق کلیدی تصاویر تمرکز کند. CBAM نیز عملکرد رقابتی از خود نشان داد و در معیار فراخوانی نتایج قابل‌توجهی داشت، در حالی که PANet به دلیل ساختار سلسله‌مراتبی خود، عملکرد ضعیف‌تری داشت. DANet از دو سازوکار توجه مجزا استفاده می‌کند: ماژول توجه موقعیتی (PAM) و ماژول توجه کانالی (CAM)، که به مدل اجازه می‌دهد هم وابستگی‌های فضایی و هم روابط کانالی را شناسایی کند. PAM روابط فضایی بین پیکسل‌ها را مدل‌سازی کرده و وزن‌های توجه را بر اساس شباهت بین پیکسل‌ها محاسبه می‌کند، در حالی که CAM بر اهمیت کانال‌های اطلاعاتی تمرکز دارد. ترکیب PAM و CAM به DANet امکان می‌دهد تا ویژگی‌های جامع‌تری ارائه دهد، که برای تصاویر هیستوپاتولوژیک حیاتی است. این سازوکار توجه دوگانه به‌ویژه در شناسایی الگوهای پیچیده و ویژگی‌های خاص کانالی مؤثر است. همچنین، DANet نقشه‌های ویژگی را به‌طور تطبیقی اصلاح می‌کند و توانایی تمرکز بر مناطق تشخیصی مرتبط را افزایش می‌دهد. در مقایسه، CBAM توجه را به‌صورت متوالی اعمال می‌کند و PANet بر تجمیع ویژگی‌های چندمقیاسی تمرکز دارد، اما فاقد سازوکارهای توجه صریح DANet است. تصاویر هیستوپاتولوژیک معمولاً پیچیدگی‌های زیادی دارند، که DANet به خوبی با آن‌ها کنار می‌آید. در نتیجه، عملکرد برتر DANet ناشی از سازوکار توجه دوگانه‌اش است که به‌طور مؤثر

وابستگی‌های فضایی و کانالی را شناسایی می‌کند. این تحقیق نه‌تنها برتری سازوکارهای توجه دوگانه را تأیید می‌کند، بلکه راه‌حلی عملی برای کاهش خطاهای تشخیصی و حمایت از تصمیم‌گیری‌های بالینی ارائه می‌دهد. در آینده، بررسی نقشه‌های توجه و توسعه سازوکارهای نوآورانه می‌تواند به بهبود بیشتر عملکرد مدل‌ها منجر شود.

مراجع

1. Giaquinto, AN., Sung, H., Newman, LA., Freedman, RA., Smith, RA., Star, J., et al. 2024. Breast cancer statistics. CA: a cancer journal for clinicians. 74(6):477-95.
2. Mahmood, T., Li, J., Pei, Y., Akhtar, F., Imran, A., Rehman, KU. 2020. A brief survey on breast cancer diagnostic with deep learning schemes using multi-image modalities. IEEE Access. 8:165779-809.
3. Shahidi, F., Daud, SM., Abas, H., Ahmad, NA., Maarop, N. 2020. Breast cancer classification using deep learning approaches and histopathology image: A comparison study. Ieee Access. 8:187531-52.
4. Balkenende, L., Teuwen, J., Mann, RM. 2022. editors. Application of deep learning in breast cancer imaging. Seminars in Nuclear Medicine; Elsevier.
5. Li, X., Li, M., Yan, P., Li, G., Jiang, Y., Luo, H. et al. 2023. Deep learning attention mechanism in medical image analysis: Basics and beyonds. International Journal of Network Dynamics and Intelligence. 93- 116.
6. Ding, R., Zhou, X., Tan, D., Su, Y., Jiang, C., Yu, G. et al. 2024. A deep multi-branch attention model for histopathological breast cancer image classification. Complex & Intelligent Systems. 10(3):4571-87.
7. Souza, MD., Ananth Prabhu, G., Kumara, V. 2025. Advanced Breast Cancer Detection Using Spatial Attention and Neural Architecture Search (SANAS-Net). SN Computer Science. 6(1):1-12.
8. Anari, S., Sadeghi, S., Sheikhi, G., Ranjbarzadeh, R., Bendeche, M. 2025. Explainable attention based breast tumor segmentation using a combination of UNet, ResNet, DenseNet, and EfficientNet models. Scientific Reports. 15(1):1027.
9. Araújo, T., Aresta, G., Castro, E., Rouco, J., Aguiar, P., Eloy, C. et al. 2017. Classification of breast cancer histology images using convolutional neural networks. PloS one. 12(6):e0177544.
10. Simonyan, EO., Badejo, JA., Weijin, JS. 2024. Histopathological breast cancer classification using CNN. Materials Today: Proceedings. 105:268-75.
11. Wang, K., Liew, JH., Zou, Y., Zhou, D., Feng, J. 2019. editors. Panet: Few-shot image semantic segmentation with prototype alignment. proceedings of the IEEE/CVF international conference on computer vision.

- proach to hyperparameter optimization. *Journal of Machine Learning Research*. 18(185):1-52.
12. Fu, J., Liu, J., Tian, H., Li, Y., Bao, Y., Fang, Z. et al., 2019. editors. Dual attention network for scene segmentation. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*.
 13. Toğaçar, M., Özkurt, KB., Ergen, B., Cömert, Z. 2020. BreastNet: A novel convolutional neural network model through histopathological images for the diagnosis of breast cancer. *Physica A: Statistical Mechanics and its Applications*. 545:123592.
 14. Zou, Y., Zhang, J., Huang, S., Liu, B. 2022. Breast cancer histopathological image classification using attention high-order deep network. *International Journal of Imaging Systems and Technology*. 32(1):266-79.
 15. Toa, CK., Elsayed, M., Sim, KS. 2024. Deep residual learning with attention mechanism for breast cancer classification. *Soft Computing*. 28(15):9025-35.
 16. Li, X., Shen, X., Zhou, Y., Wang, X., Li, T-Q. 2020. Classification of breast cancer histopathological images using interleaved DenseNet with SENet (IDSNet). *PloS one*. 15(5):e0232127.
 17. Guo, C., Zhou, Q., Jiao, J., Li, Q., Zhu, L. 2024. A Modified MobileNetV3 Model Using an Attention Mechanism for Eight-Class Classification of Breast Cancer Pathological Images. *Applied Sciences*. 14(17):7564.
 18. Gautam, A., Singh, SK., editors. 2023. Analysis of Deep Learning Models to Detect Breast Cancer from Histopathology Images. *TENCON 2023-2023 IEEE Region 10 Conference (TENCON)*; IEEE.
 19. Maleki, A., Raahemi, M., Nasiri, H. 2023. Breast cancer diagnosis from histopathology images using deep neural network and XGBoost. *Biomedical Signal Processing and Control*. 86:105152.
 20. Roy, S., Jain, PK., Tadepalli, K., Reddy, BP. 2024. Forward attention-based deep network for classification of breast histopathology image. *Multimedia Tools and Applications*. 1-30.
 21. Abdulaal, AH., Valizadeh, M., Albaker, BM., Yassin, RA., Amirani, MC., Shah, AS. 2024. Enhancing Breast Cancer Classification using a Modified GoogLeNet Architecture with Attention Mechanism. *Al-Iraqia Journal for Scientific Engineering Research*. 3(1):47-63.
 22. Spanhol, FA., Oliveira, LS., Petitjean, C., Heutte, L., editors. 2016. Breast cancer histopathological image classification using convolutional neural networks. 2016 international joint conference on neural networks (IJCNN); IEEE.
 23. Wahab, N., Khan, A., Lee, YS. 2019. Transfer learning based deep CNN for segmentation and detection of mitoses in breast cancer histopathological images. *Microscopy*. 68(3):216-33.
 24. Davis, J., Goadrich, M., editors. 2006. The relationship between Precision-Recall and ROC curves. *Proceedings of the 23rd international conference on Machine learning*.
 25. Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A., Talwalkar, A. 2018. Hyperband: A novel bandit-based ap-