

ارتقاء بازیابی تصاویر پزشکی از طریق یادگیری عمیق کد درهم سازی با استفاده از بهبود روش طبقه بندی و داده ها

رضا شمسایی^۱، نسیم نظام دوست^۲، پروانه خدائی^۳

- ۱- عضو هیئت علمی، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی سجاد، مشهد
- ۲- دانشجوی کارشناسی ارشد مهندسی کامپیوتر هوش مصنوعی و رباتیک، دانشگاه صنعتی سجاد، مشهد
- ۳- دانشجوی کارشناسی ارشد مهندسی کامپیوتر هوش مصنوعی و رباتیک، دانشگاه صنعتی سجاد، مشهد

چکیده

استفاده از یادگیری عمیق برای تکنیک های بازیابی تصاویر و تفسیر داده های پزشکی مربوط به تومور نتایج امید بخشی از خود ارایه نموده اند و در تشخیص و ارائه راه درمانی مناسب؛ از اهمیت زیادی برخوردار هستند. با این حال چالش های زیادی برای رادیولوژیست ها در غربالگری بالینی و تشخیص وجود دارد که بر عملکرد بازیابی تصاویر پزشکی تاثیر زیادی دارد. در این مقاله روشی بهبود یافته با نام Y-Net بهبود یافته، در جهت جلوگیری از این گونه چالش ها و ابهامات در بازیابی نمونه های پزشکی ارائه شده است، شبکه Y-Net از مجموعه ای از لایه های پیچشی تشکیل شده که به طور موازی برای استخراج ویژگی های پیچیده تصاویر پزشکی عمل می کنند. ما در Y-Net ارتقاء یافته در بخش طبقه بندی، از یک طبقه بند جدید برای کلاسه بندی داده ها استفاده کرده ایم و روش طبقه بند با عنوان Attention classifier جایگزین روش قبلی کرده ایم. آزمایش ها بر روی مجموعه داده های (JSRT) نشان دهنده آنست که Y-Net ارتقاء یافته، می تواند نسبت به Y-Net پایه به دلیل تمرکز بیشتر و بهتر بر روی داده ها در مرحله کلاسه بندی، ابهام مناطق غیرطبیعی پاتولوژیک را کاهش دهد. عملکرد بازیابی روش پیشنهادی با سایر الگوریتم های رقیب، براساس $map@5$ ، $map@10$ ، $map@20$ و $map@50$ اندازه گیری شده است که نشان دهند افزایشهایی ۰/۵۲٪، ۰/۴۵٪، ۰/۳۴٪، و ۰/۲۵٪ نسبت به روش قبلی است.

کلمات کلیدی: بازیابی تصاویر پزشکی، روش های درهم عمیق، تصاویر مبتنی بر محتوا، ویژگی های کانولوشنال

مقدمه

بازیابی تصاویر پزشکی [۱] یک حوزه مهم در حوزه هوش مصنوعی و بینایی ماشین است که به کمک الگوریتم ها و فنون مختلف، ما را قادر می سازد تا از طریق تحلیل و استخراج اطلاعات از تصاویر پزشکی، به شناسایی و تشخیص بیماری ها، تشخیص مشکلات سلامتی و ارائه راهکارهای بهبودی برای بیماران کمک کنیم. پیشرفت فناوری و رشد روزافزون در حوزه تصویربرداری پزشکی، بازیابی تصویر مبتنی بر محتوا (CBIR) به عنوان یکی از مهمترین و چالش برانگیزترین مسائل در این زمینه مطرح شده است.

در حوزه پزشکی، تصاویر پزشکی از جمله عکس های رادیولوژی (مانند اشعه ایکس، سونوگرافی، اسکن توموگرافی گسیل پوزیترون (PET)، تصویربرداری تشدید مغناطیسی (MRI)) و تصاویر بالینی (مانند عکس های میکروسکوپی، تصاویر جراحی و

تصاویر ارتوپدی) شامل اطلاعات بسیار زیادی درباره وضعیت بیماری و سلامتی فرد را در بر می گیرند. بنابراین، مدیریت چنین پایگاه داده های عظیمی به طور کارآمد و سازنده یک نیاز حیاتی است.

با توجه به حجم بزرگ تصاویر پزشکی که در پایگاه داده های بیمارستان ها و مراکز پزشکی ذخیره می شوند، استفاده از روش های بازایی تصاویر پزشکی به عنوان یک ابزار مفید واجد اهمیت است. این روش ها می توانند زمان و هزینه های مربوط به تشخیص و درمان را به میزان قابل توجهی کاهش دهند و به پزشکان کمک کنند تا تصاویر پزشکی را به صورت مؤثرتر و سریع تر تحلیل و تفسیر کنند.

با این حال، با توجه به پیشرفت های اخیر در حوزه هوش مصنوعی و شبکه های عصبی عمیق، بازایی تصاویر پزشکی نمی تواند بهبود واضحی در دقت و کارایی داشته باشد. بازایی سطح نمونه در حوزه پزشکی چالش های خاصی را به همراه دارد. برای مثال، بیماری های مختلف می توانند نشانه های غیرطبیعی مشابهی داشته باشند و نشانه های غیرطبیعی مختلف ممکن است در مراحل مختلف یک بیماری مشاهده شوند. به عبارت دیگر، تشخیص دقیق و قابل اعتماد بر اساس بازایی سطح نمونه در حوزه پزشکی همواره چالش هایی را در برداشته است.

الگوریتم ها و مدل های عمیق مبتنی بر یادگیری عمیق، مانند شبکه های عصبی پیچشی (CNN)، [2] می توانند ویژگی های پرباری را از تصاویر پزشکی استخراج کنند و با استفاده از روش های مقایسه و مطابقت، تصاویر مشابه را پیدا کنند. یادگیری عمیق به عنوان یک روش قدرتمند در حوزه بینایی کامپیوتر و پردازش تصویر، توانایی برتری در تشخیص تصاویر پزشکی نشان داده است. با استفاده از الگوریتم های یادگیری عمیق، رایانه ها قادرند الگوها و ویژگی های مهم را به طور خودکار از داده ها استخراج کرده و اطلاعات به روز و قابل استفاده را به پزشکان ارائه دهند. این قابلیت ها باعث شده است که یادگیری عمیق به عنوان یک ابزار قدرتمند در تشخیص و تفسیر تصاویر پزشکی مورد توجه قرار گیرد.

با افزایش حجم تصاویر پزشکی در سیستم های ذخیره سازی و انتشار، نیاز به روش های مؤثرتر برای مدیریت و بازایی این تصاویر اهمیت بیشتری پیدا کرده است. یکی از ابزارهای مفید برای بازایی مؤثر و سریع تصاویر پزشکی، استفاده از کدهای درهم است.

کدهای درهم [3] در این حوزه به تصاویر پزشکی اختصاص داده می شوند تا هویت و یکتایی تصاویر را تضمین کنند و به عنوان امضای دیجیتالی برای تصاویر عمل کنند. با استفاده از تابع درهم، می توان یک مقدار درهم یکتا برای هر تصویر پزشکی محاسبه کرد. این مقدار درهم به عنوان یک شناسه منحصر به فرد برای تصویر استفاده می شود.

بازایی تصاویر پزشکی با استفاده از کدهای درهم به تعداد زیادی مزیت می رسد. ابتدا، با استفاده از مقدار درهم، می توان تصاویر را به سرعت و به صورت یکتا در سیستم ذخیره و مدیریت کرد. هرگاه نیاز به بازایی یک تصویر پزشکی وجود داشت، می توان با مقایسه مقدار درهم آن با مقادیر کد درهم ذخیره شده، تصویر مورد نظر را به سرعت و دقت پیدا کرد. در ادامه، با استفاده از کدهای درهم، می توان از امنیت بالایی در مدیریت تصاویر پزشکی برخوردار شد. با محاسبه مقدار کد درهم برای هر تصویر، می توان اطمینان حاصل کرد که تصویر اصلی تغییر نکرده است. در صورتی که تصویر دچار تغییری شود (مثلاً با ویرایش یا تغییرات غیرمجاز در داده ها)، مقدار کد درهم تغییر خواهد کرد و این تغییر قابل مشاهده است. این به پزشکان و متخصصان پزشکی امکان می دهد تا از اصالت و صحت تصاویر اطمینان حاصل کنند و در تشخیص و درمان بیماران اعتماد بیشتری داشته باشند.

استفاده از کدهای درهم در بازایی تصاویر پزشکی یک روش مؤثر و قابل اعتماد است که به طبقه بندی، مدیریت و

امنیت تصاویر کمک می‌کند. با ایجاد ارتباط بین مقادیر کد درهم و تصاویر پزشکی، می‌توان به سرعت تصاویر را جستجو و بازیابی کرد و در عین حال از تغییرات غیرمجاز در تصاویر جلوگیری کرد.

در نتیجه، بازیابی تصاویر پزشکی با توجه به اهمیت و کاربردهای آن، همچنان یک حوزه فعال پژوهشی است که نیازمند تلاش‌های بیشتر در توسعه روش‌ها، بهبود کیفیت و دقت، و استفاده بهینه از فناوری‌های جدید است تا بتواند به صورت گسترده‌تر در عمل مورد استفاده قرار بگیرد و به بهبود سلامت و درمان بیماران کمک کند.

در این مقاله، بهبود بازیابی تصاویر پزشکی با استفاده از یادگیری عمیق و کدهای درهم‌سازی در مقاله پایه را مورد بررسی قرار می‌دهیم. ولی هدف اصلی این تحقیق، ارتقای روش طبقه‌بندی داده‌ها در بازیابی تصاویر پزشکی است. برای این منظور، یک چارچوب جدید به نام Y-Net را ارائه می‌دهیم که تصاویر را به کدهای درهم‌سازی تبدیل می‌کند. این چارچوب با استفاده از تجمیع ویژگی‌های کانولوشنی و رمزگذاری، توانایی استخراج اطلاعات مفید و مرتبط از تصاویر را دارا می‌باشد. علاوه بر این، در این تحقیق، روشی جدید برای کلاسه‌بندی داده‌ها در بخش طبقه‌بندی نیز ارائه می‌شود. این روش طبقه‌بند جدید قادر است بهبود قابل توجهی در کاهش ابهام مناطق غیرطبیعی و پاتولوژیک داشته باشد و عملکرد بازیابی تصاویر را بهبود بخشد و دقت آن را افزایش دهد.

با توجه به اهمیت بازیابی تصاویر پزشکی در تشخیص و درمان بیماری‌ها، تحقیقات انجام شده در این زمینه می‌تواند بهبود قابل توجهی در عملکرد سیستم‌های تشخیصی و تصویربرداری پزشکی داشته باشد. امیدواریم که این تحقیق، گامی مهم در جهت بهبود بازیابی تصاویر پزشکی و کمک به متخصصان در تفسیر و تشخیص صحیح تصاویر باشد.

بقیه این مقاله به شرح زیر سازماندهی شده است: بخش ۲ کارهای مرتبط را مورد بحث قرار می‌دهیم. بخش ۳ به طور گسترده روش پیشنهادی را بر روی یک مجموعه داده تصاویر پزشکی ارزیابی می‌کنیم. بخش ۴ ارزیابی و مقایسه روش پیشنهادی مورد بررسی قرار می‌گیرد و بخش ۵ نکات پایانی و نتیجه‌گیری را بیان می‌کنیم.

۲- مروری بر کارهای گذشته

به دلیل اهمیت موضوع بازیابی تصاویر پزشکی تحقیقات زیادی در این حوزه صورت گرفته و در حال انجام است که در ادامه بر اساس تاریخ ارائه به آن می‌پردازیم. یادگیری عمیق موفقیت قابل قبولی را در وظایف تشخیص تصویر نشان داده است. با افزایش حجم داده‌ها، یادگیری عمیق به رایانه‌ها اجازه می‌دهد تا الگوهایی را به طور خودکار در داده‌ها یاد بگیرند که می‌تواند برای پیش‌بینی و تشخیص مفید باشد.

روش دقت منطقه‌ای^۱ [4] به بهبود عملکرد جستجوی تصاویر با استفاده از استخراج و استفاده از ویژگی‌های عمیق تصاویر به منظور جستجوی دقیق تصاویر می‌پردازد. این روش از یک شبکه عصبی عمیق برای استخراج ویژگی‌های تصاویر استفاده می‌کند و سپس با توجه به مناطق مختلف تصویر، ویژگی‌های مکانی و معنایی را تا حد زیادی در نظر می‌گیرد. این کار باعث افزایش دقت جستجو تصاویر در مواردی می‌شود که ویژگی‌های مناطق مختلف تصویر به طور متفاوت ارتباط دارد.

¹ Regional Attention

میانگین دقت متوسط^۱ به عنوان یک پروتکل ارزیابی در این مقاله استفاده شده است.

یکی دیگر از روش‌های انجام‌شده در این زمینه، تقسیم‌بندی ناهمگن کاملاً خودکار با استفاده از دستگاه بردار پشتیبانی (FAHS- SVM) می‌باشد که بیشتر در تشخیص تومور مغزی استفاده شده است. تصویربرداری رزونانس مغناطیسی^۲ یک ابزار تشخیصی برای مطالعه و آزمایش آناتومی انسان است. افزایش عروق تومور منجر به جذب ترجیحی عامل کنتراست می‌شود و می‌توان برای مشاهده تومورهای بافت طبیعی اطراف آن‌ها بهتر استفاده کرد که تمایز بین بیماری‌های بدخیم یا خوش‌خیم را افزایش دهد. در این روش بعد از پیش‌برازش داده‌ها^۳، تمام بافت‌های غیر مغزی در تصویربرداری مغز حذف می‌شوند و سپس در مرحله مورفولوژی، تصویر رزونانس مغناطیسی مغز از پیش‌پردازش شده به یک تصویر مضاعف با آستانه ۱۲۸ برای برش تبدیل می‌شود. مقادیر پیکسل بالاتر از آستانه‌های مشخص شده به صورت سفید و سایر مناطق به عنوان سیاه مشخص می‌شوند. این دو اجازه می‌دهد مناطق مختلفی در اطراف این بیماری ایجاد شود. در مرحله دوم، از فرسایش مورفولوژی برای استخراج پیکسل‌های سفید استفاده می‌شود. در نهایت، ناحیه فرسوده و تصویر اصلی به دو ناحیه مساوی تقسیم می‌شوند و ناحیه‌ای با پیکسل‌های سیاه ناشی از فرسایش به عنوان ماسکی از تصویر رزونانس مغناطیسی مغز در نظر گرفته می‌شود. بافت‌های دارای ساختار مغزی پیچیده هستند و ویژگی‌های شدت آن‌ها برای تقسیم‌بندی دقیق تومور کافی نیست. برای بهبود دقت تقسیم‌بندی، از ویژگی‌های بافت استفاده می‌شود. در این مقاله ویژگی‌های بافتی بر اساس تحلیل بافت ارزیابی می‌شوند. Textons عناصر کوچک تصویری هستند که توسط یک تصویر متحرک با یک مجموعه فیلتر خاص ایجاد می‌شوند. دقت تقسیم‌بندی در مقایسه با سایر روش‌های موجود زیاد است. [5]

یکی از روش‌های دیگر این است که کاربر برخی از تصاویر را به عنوان بازخورد مثبت و تعدادی از تصاویر ناچیز را به عنوان بازخورد منفی علامت‌گذاری یا برچسب‌گذاری کند. بر اساس این نمونه‌های برچسب زده شده، سیستم^۴ CBMIR سپس فرآیند ارزیابی خود را اصلاح می‌کند. این ایده بر اساس تصاویر برتر ارزیابی شده با استفاده از گروه هشت معیار یا ضرایب تشابه پذیرفته شده است. سپس، تصاویر برتر ارزیابی شده مشابه به عنوان رأی‌دهنده برای انتخاب بهترین عملکرد برای هر کلاس تصویر استفاده می‌شود. گروهی از ویژگی‌های رنگ و بافت با استفاده از روش‌های معروف استخراج، بر اساس رنگ و ویژگی‌های بافت، استخراج می‌شود. برای اندازه‌گیری‌های شباهت، از هشت ضریب تشابه رایج به عنوان مبنای تشابه استفاده شده است. آزمایش با مجموعه داده‌های پزشکی مشهور Kvasir نشان می‌دهد که این روش یک روش بسیار ساده برای افزایش و بهبود اثربخشی ارزیابی تصاویر پزشکی مرتبط ارائه می‌دهد. [6]

یکی دیگر از روش‌های ارزیابی تصاویر، استفاده از اطلاعات مرتبه دوم^۵ است. [7] اطلاعات مرتبه دوم در بینایی کامپیوتر توجه فزاینده‌ای را به خود جلب کرده و می‌توانند در ارزیابی تصویر به شکل همبستگی خودکار فضایی ویژگی‌ها یا با شباهت‌های مرتبه دوم در یک فضای متریک، مورد بهره‌برداری قرار گیرند. در این مقاله از روابط مرتبه دوم بین ویژگی‌ها در مکان‌های فضایی مختلف بهره‌برداری می‌کند و با شباهت توصیف‌گر مرتبه دوم ترکیب می‌کند تا توصیفگرهای ویژگی را برای ارزیابی و تطبیق تصویر بهبود دهد. در اینجا ابتدا تجمیع میانگین تعمیم یافته^۶ را ارائه می‌کند که با ادغام مرتبه دوم خود گسترش می‌یابد، به دنبال آن از دست دادن شباهت مرتبه دوم و نرمال سازی توصیفگر است. از یک تصویر ورودی I که از

¹ Mean average precision

² MRI

³ preprocessing

⁴ Content Based Medical Image Retrieval

⁵ econd-order information

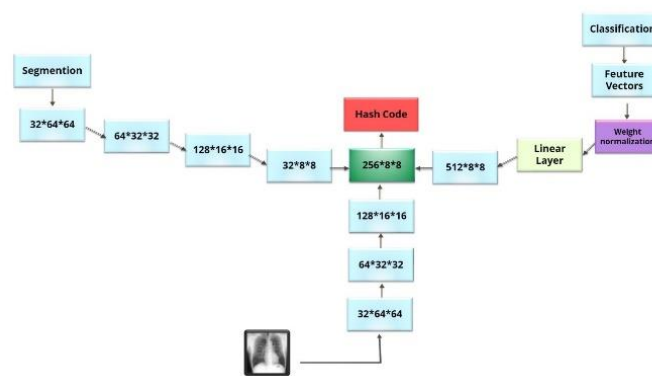
⁶ Generalised-Mean pooling

طریق یک شبکه کاملاً کانولوشنل پردازش شده است که با θ نشان داده شده است، یک نقشه ویژگی $f = \theta(I) \in R^{h,w,d}$ به دست می آوریم که در آن w, h و d ارتفاع، عرض و ابعاد ویژگی هستند. برای $h, w > 1$ ، ادغام میانگین تعمیم یافته (GeM) به عنوان یک روش منعطف برای تجمیع نقشه ویژگی در یک بردار توصیفگر واحد $D = GeM(f, P)$ پیشنهاد شده است. ادغام GeM با پارامتر قابل یادگیری p به صورت زیر تعریف شده است.

$$GeM(f, p) = \left(\frac{1}{N} \sum_{i=0}^N f_i^p \right)^{\frac{1}{p}} \quad (1)$$

روش DDMH¹ [8] یک چارچوب درهم سازی تکانه عمیق (DDMH) برای جستجوی تصویر عصبی پیشنهاد می کند. به طور خاص، ابتدا تلفات سه گانه اصلی را بررسی و متوجه می شود که این تابع از دست دادن را می توان با حاصلضرب داخلی جفت کد درهم تعیین کرد. بر این اساس، هنجارهای کد درهم و جهت های کد درهم را از هم جدا کرده و نقش هر بخش را تحلیل میکنند. با جدا کردن تابع از دست دادن از هنجار کد درهم، یک تلفات سه گانه جدا شده منحصر به فرد را پیشنهاد می کند، که می تواند به طور موثر جفت های نمونه مثبت و منفی را با اختلاف فاصله همینگ² مورد نظر برای کدهای درهم با طول های مختلف انتشار دهد. در این مقاله با از دست دادن سه گانه پیشنهادی و استراتژی سه گانه حرکت، یک چارچوب درهم عمیق قابل آموزش سرتاسری برای جستجوی تصویر عصبی طراحی می کند. در این مقاله مقادیر میانگین دقت متوسط برای ارزیابی عملکرد روش پیشنهادی استفاده شده است.

۳- روش پیشنهادی



شکل ۱: دیاگرام بالا روش مقاله پایه را نمایش می دهد. در این شکل قسمت هایی که با بلاک های آبی رنگ مشخص شده اند با استفاده از شبکه عصبی سریع مبتنی بر منطقه عملیات آشکارسازی شامل کلاسه بندی و مکان یابی را انجام می دهد. بلاک سبز رنگ نشان دهنده ادغام دو لایه کانولوشنل می باشد که نتیجه این ادغام رسیدن به لایه ای با ویژگی های برتر می باشد. شاخه اصلی شامل قسمت از پایین به بالا بعد از گرفتن تصویر

¹ Deep Disentangled Hashing with Momentum

² Hamming distance

ورودی میباشد که تصویر را به لایه های کانولوشنالی تبدیل میکند.

در این مقاله یک چارچوب عمیق داده شده در مقاله پایه [۹] را با بررسی شاخه های موجود در آن و انتخاب و پیدا کردن روشهای جدید تر مورد بهبود قرار داده ایم. این بهبود با ارتقا روش دسته بندی داده ها انجام شده است که در ادامه به توضیح آن میپردازیم.

همانطور که در شکل ۱ میبینیم Y-Net می تواند ویژگی های کانولوشنالی^۱ را با ادغام دو شاخه طبقه بندی و کلاسه بندی بیاموزد و می تواند به طور موثر ویژگی های متمایز را در کدهای درهم در مرحله بازیابی جاسازی کند. هدف Y-Net، ایجاد کدهای درهم بسیار متمایز از ویژگی های پیچشی آموخته شده می باشد. Y-NET شامل سه بخش است، شاخه اصلی، شاخه R-MAC (شاخه راست) و شاخه FPN^۲ (شاخه چپ). در مرحله آموزش یک تصویر را وارد شاخه اصلی می کند و به سمت گره هسته به سمت جلو تغذیه می کند. گره هسته یک لایه کانولوشن است و به دنبال آن شاخه R-MAC و شاخه FPN قرار می گیرد. در شاخه R-MAC، تلفات طبقه بندی فاصله درون کلاسی را به حداقل می رساند و فاصله بین طبقاتی را به حداکثر می رساند. که در این قسمت ما برای بهبود از تغییرات روش طبقه بندی داده ها استفاده کردیم و شاخه FPN می تواند تفاوت های درون کلاسی را با آموزش تقسیم بندی پیکسلی تعیین کند تا کاهش فاصله درون کلاسی در شاخه R-MAC را متعادل کند.

۳-۱ شاخه اصلی

در مرحله اول تصویر ورودی با وضوح $۲۵۶ \times ۲۵۶ \times ۳$ به شاخه اصلی فید فوروارد می شود. شاخه اصلی یک سلسله مراتب ویژگی متشکل از یک بلوک از پایین به بالا در سه مقیاس را محاسبه می کند. در هر مقیاس، ما از خروجی فعال سازی ویژگی بلوک پایین به بالا برای بدست آوردن یک خروجی با ویژگی های بهتر استفاده می کنیم. سه بلوک پایین به بالا با افزودن در شاخه FPN ادغام می شوند.

در گره هسته شاخه اصلی، نقشه های ویژگی کانولوشنالی $X \in \mathbb{R}^{C \times H \times W}$ را می توان در یک تانسور به اندازه $C \times H \times W$ مرتب کرد، که در آن H و W ارتفاع و عرض هر نقشه ویژگی را نشان می دهند. C تعداد نقشه های ویژگی (یا کانال ها) در لایه کانولوشن را نشان می دهد. نقشه های ویژگی کانولوشن X بیشتر به شاخه R-MAC و شاخه FPN ارسال می شوند. ما تجمیع ویژگی را برای تولید یک کد درهم k بیت از نقشه های ویژگی کانولوشنالی آموخته شده X با $C \times H \times W$ در گره هسته اعمال می کنیم.

۳-۲ شاخه R-MAC

شاخه R-MAC (شاخه راست Y-NET) برای آموزش طبقه بندی مورد استفاده قرار می گیرد. [10] اطلاعات معنایی آگاه از کلاس مناطق آسیب شناسی از شاخه R-MAC موثر در نقشه های ویژگی کانولوشن در گره هسته آموخته می شود. شاخه R-MAC شامل یک لایه کانولوشن است که در ادامه داده ها در یک مدل طبقه بند نرم^۳ کلاسه بندی میشوند و سپس یک بلوک R-MAC یک بردار ویژگی ایجاد می کند. بردار ویژگی پس از نرمال سازی به یک لایه خطی نگاشت می شود. در روش Y-NET اصلی، برای دسته بندی داده ها در شاخه R-MAC از یک مدل دسته بندی به صورت زیر استفاده میکنند [11]:

¹ convolutional features

² Feature Pyramid Networks

³ Softmax Classifier

$$b + x \cdot w = y \quad (2)$$

که در اینجا x بردار ورودی با ابعاد 512 (ویژگی‌های استخراج شده) ، w ماتریس وزن‌های لایه با ابعاد (تعداد کلاس‌ها 1×512) و b بردار بایاس با ابعاد تعداد کلاس‌ها و y خروجی نهایی با ابعاد تعداد کلاس‌ها که به تعداد کلاس‌ها احتمال تعلق اختصاص می‌دهد.

با توجه به اینکه داده‌های ما تصاویر نامتوازن و پیچیده‌ای هستند، انتخاب روش طبقه‌بند توجه^۲ باعث افزایش تمرکز بر روی ویژگی‌های مهم تصاویر و همچنین کاهش تاثیر نویز و در نتیجه بهبود عملکرد دسته‌بندی داده‌ها میشود [۱۲]. در ادامه به معرفی روش طبقه‌بندی گفته شده می‌پردازیم.

۱-۲-۳ روش طبقه‌بند توجه

مکانیزم طبقه‌بند توجه در طبقه‌بندی به این صورت عمل می‌کند که به مدل اجازه می‌دهد روی بخش‌های مهم‌تر ورودی‌ها تمرکز کند. به طور ساده‌تر، این روش به مدل می‌گوید که کدام قسمت از داده (مثلاً ویژگی‌ها یا نواحی تصویر) برای پیش‌بینی نهایی مهم‌تر است. عملکرد روش به این صورت است که:

- ابتدا وزن‌ها محاسبه می‌شوند: هر ویژگی یک اهمیت یا وزن دریافت می‌کند.
- سپس ویژگی‌ها با این وزن‌ها ترکیب می‌شوند و ویژگی‌های مهم‌تر تقویت می‌شوند.
- در نهایت، این ویژگی‌های وزن‌دار از طریق یک لایه طبقه‌بندی می‌گذرند تا کلاس پیش‌بینی شود:
خروجی = لایه خطی

به طور خلاصه، طبقه‌بند توجه و طبقه‌بند نرم هر دو در دسته‌بندی داده‌ها عملکرد قابل قبولی دارند طبقه‌بند توجه اجازه می‌دهد که مدل روی نواحی مهم‌تر تمرکز کند و در نتیجه عملکرد بهتری در دقت پیش‌بینی بهتری در دسته‌بندی تصاویر پزشکی ارائه می‌دهد [13]. این رویکرد به خصوص در داده‌های پیچیده‌ای مانند نمونه استفاده شده در Y-Net موثرتر است.

۳-۳ شاخه FPN

شاخه FPN (شاخه چپ Y-NET) برای آموزش تقسیم‌بندی مورد استفاده قرار می‌گیرد تفاوت‌های ظریف فضایی مناطق آسیب‌شناسی از شاخه FPN به طور موثر در نقشه‌های ویژگی کانونولوشن در گره هسته آموخته می‌شود. نقشه‌های ویژگی کانونولوشن از طریق تجمیع ویژگی در مرحله آزمایش به کدهای درهم نگاشت می‌شوند. [۱۴]

شاخه FPN که از آموزش تقسیم‌بندی سود می‌برد، تفاوت‌های ظریف چند مقیاسی مناطق آسیب‌شناسی را در مراحل مختلف بررسی می‌کند و سپس به گره اصلی در شاخه اصلی بازخورد می‌دهد. FPN از ویژگی‌های پیچشی از سطوح پایین به بالا برای استخراج اطلاعات چند مقیاسی با ویژگی‌های هرمی استفاده می‌کند. FPN قابلیت استخراج ویژگی قدرتمند خود را

¹ $classes^n$

² Attention Classifier

برای دستیابی به دقت بالاتر نشان می دهد.

۴-۳ تابع زیان

در Y-Net، طبقه بندی در شاخه R-MAC را با تقسیم بندی در شاخه FPN ادغام می کنیم تا اطلاعات معنایی و فضایی مناطق آسیب شناسی را به طور همزمان بیاموزیم. برای متعادل کردن نتیجه حاصل از ادغام دو شاخه سمت راست و چپ، مقاله پایه یک تابع زیان ۱ را طراحی کرده است تا یادگیری طبقه بندی و تقسیم بندی را یکسان کند که به صورت زیر تعریف میشود.

$$L = wL_1 + (1 - w)L_r \quad (3)$$

که در آن L_r کاهش دایره ۲ برای آموزش طبقه بندی است، L_1 نشان دهنده افت آنتروپی برای آموزش تقسیم بندی پیکسلی است، و w ضریب وزن است. بر اساس ضایعات جفت شده که اتلاف طبقه بندی و تقسیم بندی را متحد می کند، گره اصلی شاخه اصلی می تواند به طور موثر اطلاعات فضایی چند مقیاسی را از آموزش تقسیم بندی و اطلاعات معنایی آگاه از کلاس را از آموزش طبقه بندی به طور همزمان حفظ کند. با کمک تابع زیان که بین دو ضرر تعادل ایجاد می کند، نقشه های ویژگی ردیف Y-Net اثربخشی حفظ اطلاعات از شاخه R-MAC و شاخه FPN را تأیید می کند. از این رو، ویژگی های کانولوشنی آموخته شده گره هسته را می توان برای تولید کدهای درهم برای مبارزه با شواهد مبهم مناطق آسیب شناسی استفاده کرد.

۴- ارزیابی و مقایسه روش پیشنهادی

برای ارزیابی عملکرد Y-Net پیشنهادی، آزمایش های گسترده ای روی یک مجموعه داده تصویر پزشکی عمومی انجام شده است تا اثربخشی روش خود را در مبارزه با شواهد مبهم مناطق پاتولوژیک تأیید کند. Y-Net پیشنهادی تحت چارچوب PyTorch پیاده سازی می شود و آزمایش هایی روی سیستم Core i7 با ۳.۹۰ GHZ CPU @ ۳.۴۰ GHz اجرا می شوند. ما از اعتبارسنجی متقاطع ^۳ ۵ برابر برای انتخاب بهترین مدل طبقه بندی و تقسیم بندی استفاده می کنیم.

۴-۱ داده های آموزشی

داده JSRT^۵ [۱۵] شامل ۱۵۴ تصویر ندول و ۹۳ تصویر غیر ندول رادیوگرافی قفسه سینه است. داده ها به دو کلاس به عنوان خوش خیم یا بدخیم تقسیم بندی شده است. تصاویر حاشیه نویسی برای وظایف تقسیم بندی، تصاویر باینری هستند که در آنها پیکسل ها یا ۲۵۵ برای پیش زمینه یا ۰ برای پس زمینه هستند. ما ۱۳۸ تصویر شامل ۸۹ گره بدخیم و ۴۹ ندول خوش خیم را برای تشکیل یک مجموعه آموزش و ۱۶ تصویر حاوی ۱۱ ندول بدخیم و ۵ ندول خوش خیم را برای تشکیل یک مجموعه آزمایش نمونه برداری کردیم. نسبت مجموعه آموزش و مجموعه تست ۹ : ۱ است. فقط داده های کلاسه ندول در آموزش و

¹ coupled loss

² circle loss

³ K- fold cross-validation

⁴ K=5

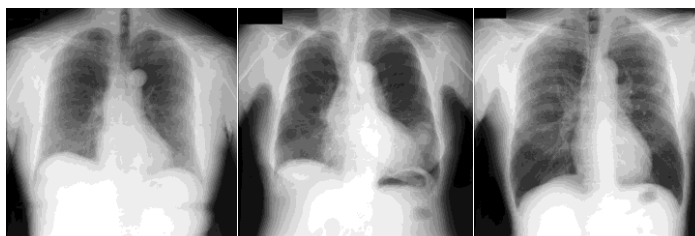
⁵ Japanese Society of Radiological Technology

آزمایش مورد بررسی قرار گرفته اند. ۹۳ تصویر غیر ندول در آموزش و آزمایش مورد استفاده قرار نگرفته اند و در جدول شماره ۱ نیز آورده نشده اند. تعداد متناقض در جدول شماره ۱ به دلیل عدم محاسبه این ۹۳ تصویر میباشد.

جدول ۱ اطلاعات و جزئیات داده JSRT شامل تصاویر ندول با دو کلاس خوش خیم و بدخیم [15]

نوع کلاسه بندی	تعداد کل تصاویر	تعداد تصاویر آموزش	تعداد تصاویر آزمایش
خوش خیم	۵۴	۴۹	۵
بد خیم	۱۰۰	۸۹	۱۱

شکل ۲ نمونه ای از تصاویر اشعه ایکس قفسه سینه است که زیر به نمایش گذاشته شده است.



شکل ۲: تصویر دیجیتال رادیوگرافی قفسه سینه با و بدون ندول ریه .

۴-۲ معیار ارزیابی

جهت ارزیابی عملکرد سیستم های بازیابی تصاویر از معیار میانگین متوسط دقت^۱ استفاده میشود که در ادامه به توضیح آن خواهیم پرداخت.

در لیست برگشتی^۲، این معیار رتبه های تصاویر مشابه تصویر پرس و جو را برای اندازه گیری کیفیت رتبه میانگین می گیرد که در اینجا عملکرد آن نسبت به لیست برگشتی ۵، ۱۰، ۲۰ و ۵۰ گزارش شده است.

این معیار معمولاً برای ارزیابی عملکرد بازیابی با فرمول زیر استفاده می شود:

$$AP = \frac{\sum_{k=1}^n P(k) \cdot rel(k)}{R} \quad (4)$$

که در اینجا R تعداد نتایج مشابه جست و جو ، P(K) نشانگر نتایج دقت بازیابی و rel(K) بیانگر یک تابع دودویی می باشد

¹ MAP

² Returned List

و n تعداد کل نتایج بازیابی شده می باشد. همه عملکردها از طریق کد درهم ۶۴ بیتی گزارش شده است.

برای بررسی عملکرد روش، رتبه‌های تصاویر مشابه تصویر پرس و جو را برای اندازه‌گیری کیفیت رتبه میانگین می‌گیرد. عمدتاً از معیارهای استاندارد بازیابی تصویر (میانگین دقت متوسط و دقت میانگین و ...) برای ارزیابی کمی استفاده میشود.

برای بررسی تشخیص صحیح نمونه‌های هر کلاس که آیا مدل توانسته تصاویر هر کلاس صحیح را بازیابی کند یا خیر؟

به طور مشترک اطلاعات معنایی آگاه از کلاس مناطق آسیب‌شناسی را بررسی می‌کنند و با بهره‌برداری از همبستگی بین برجسب‌های کلاس می‌تواند مشکل بیماری‌های مختلف که ناهنجاری‌های پاتولوژیک یکسان داشته باشند یا تظاهرات پاتولوژیک مختلف ممکن است در مراحل مختلف یک بیماری رخ دهد را در بازیابی نمونه‌های پزشکی برطرف کند.

۳-۴ نتایج مقایسه ای

برای ارزیابی عملکرد روش پیشنهادی، به مقایسه این روش با روش‌های پیشین مانند روش Regional Attention [۶] و روش SOLAR-Local [۷] که بر روی اطلاعات مکانی مرتبه دوم تمرکز می‌کند و همچنین روش DDMH [۸] می‌پردازیم.

عملکرد mAP نسبت به لیست بازگشتی ۵، ۱۰، ۲۰ و ۵۰ در مجموعه داده‌های JSRT به ترتیب در جدول ۲ گزارش شده است. و مقایسه روش‌های پیشین با روش پیشنهادی نیز در جدول ۲ مورد مقایسه قرار گرفته است. به طور کلی، وقتی لیست برگشتی طولانی می‌شود، عملکرد همه روش‌ها تا حدی کاهش می‌یابد که در روش پیشنهادی‌ها همانطور که در جدول مشاهده میشود میزان دقت در Map 50 نسبت به سایر لیست بازگشتی کمتر است اما باز هم نسبت به سایر روش‌ها در لیست بازگشتی طولانی‌تر هم بهبود عملکرد قابل ملاحظه‌ای را شاهد هستیم و Y-Net ارتقا یافته به دستاوردهای قابل توجهی از mAP نسبت به لیست برگشتی متفاوت در مجموعه داده دست می‌یابد. Y-Net ارتقا یافته همچنین بهترین عملکرد را در مجموعه داده JSRT در مقایسه با روش Y-NET و سایر روش‌های مقایسه‌ای عنوان شده به دست می‌آورد و از بین تمامی روش‌های مقایسه شده بهترین عملکرد را دارد و توانسته است بیشترین میزان کارایی بازیابی را بدست آورد.

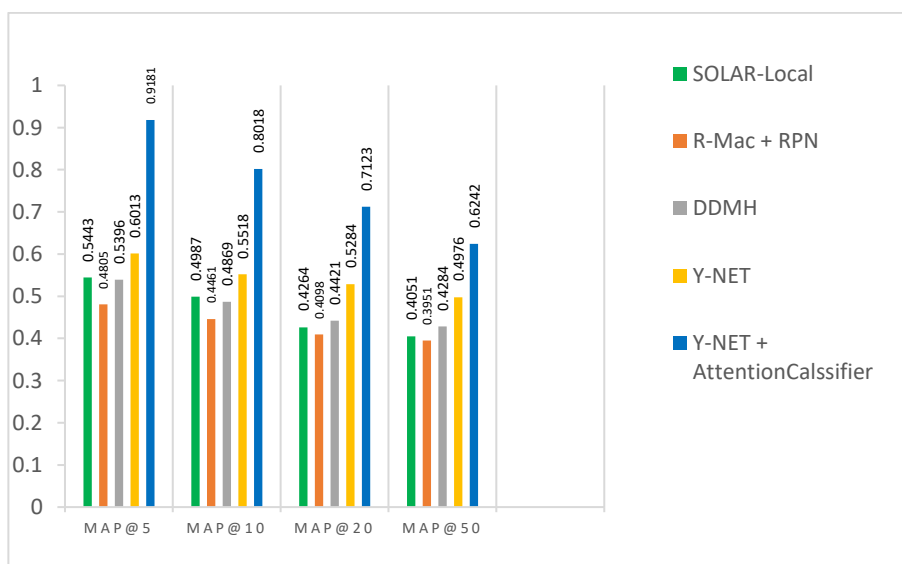
نتایج تجربی روی مجموعه داده نشان می‌دهد که Y-Net ارتقا یافته نسبت به تعداد متفاوت لیست برگشتی، از تمام روش‌های پیشین در جدول (۲) بهتر عمل می‌کند و بهترین نتایج را در مجموعه داده‌های JSRT در مقایسه با روش‌های دیگر بدست می‌آورد. تنها چالش روش ما در این است که استفاده از طبقه بند جدید پیچیدگی مدل را بیشتر کرده و محاسباتی بالاتری را نیاز دارد.

جدول ۲ عملکرد و مقایسه mAP نسبت به لیست بازگشتی ۵، ۱۰، ۲۰ و ۵۰ در مجموعه داده های JSRT

نام روش	JSRT			
	MAP@5	MAP@10	MAP@20	MAP@50
SOLARLocal	0.5443	0.4987	0.4264	0.4051
RMac+RPN	0.4805	0.4461	0.4098	0.3951
DDMH	0.5396	0.4869	0.4421	0.4284
Y-NET	0.6013	0.5518	0.5284	0.4976
+Y-NET AttentionCalssifier	0.9181	0.8018	0.7123	0.6242

نمودار ۱

عملکرد و مقایسه mAP نسبت به لیست بازگشتی ۵، ۱۰، ۲۰ و ۵۰ در مجموعه داده های JSRT



در مرحله بعد، ما می‌خواهیم اثر تابع ضرر را در آموزش شاخه R-MAC مورد بحث قرار دهیم. در روش پایه تابع ضرر جفت شده می‌تواند عملکرد بازیابی را در mAP نسبت به تعداد متغیر لیست برگشتی در مجموعه داده های JSRT بهبود بخشد. و، همانطور که در جدول ۳ نشان داده شده است، در Y-NET ارتقا یافته نیز، عدم استفاده از تابع ضرر جفت شده می‌تواند عملکرد بازیابی را در mAP نسبت به تعداد متغیر لیست برگشتی در مجموعه داده های JSRT ضعیف تر کند. البته تحقیقات ما نشان می‌دهد که روش ما با کاهش فاصله درون کلاسی و بهبود استخراج ویژگی‌ها از تصاویر و بدون استفاده از تابع ضرر، همچنان عملکرد بهتری نسبت به روش های قبلی خود دارد که در جدول ۳ نتایج را بررسی کرده ایم.

جدول ۳ عملکرد و مقایسه mAP نسبت به استفاده و عدم استفاده از تابع ضرر در شاخه R-MAC در مجموعه داده های JSRT

شاخه	JSRT			
	MAP@5	MAP@10	MAP@20	MAP@50
Y-NET	0.6013	0.5518	0.5284	0.4976
+Y-NET AttentionCalssifier withoutCircle loss	0.9181	0.8018	0.7123	0.6242
+Y-NET AttentionCalssifier	۰.۷۶۲۱	۰.۶۶۳۲	۰.۵۹۴۳	۰.۴۹۸۱

۴- نتیجه گیری

در این مقاله ما یک روش جدید پیشنهاد می‌کنیم که تصاویر را در کدهای درهم فشرده که از ویژگی‌های کانولوشن ادغام می‌شوند، رمزگذاری می‌کند. روش پیشنهادی شامل شاخه اصلی، شاخه R-MAC، شاخه FPN است. همانطور که گفته شد در این مقاله نشان دادیم که با اعمال بهبود روش طبقه بندی در قسمت R-Mac نتایج بهتری مورد انتظار است. در قسمت R-Mac (بلاک آبی پایین) ما از یک روش طبقه بندی برای دسته بندی داده ها برای کاهش فاصله درون کلاسی استفاده می‌کنیم که روش استفاده شده در این قسمت در این مقاله بهبود یافته است. هدف اولیه در این تحقیق استفاده از یک روش به روز تر در دسته بندی داده ها می باشد که انتخاب روش طبقه بند توجه است^۱ که روش پیشنهادی ما در مقایسه با سایر روش‌ها به دلیل بهره‌گیری از طبقه بند جدید و استفاده از ساختار Y-Net توانسته است با کاهش خطای دسته‌بندی در داده‌های پیچیده مانند تصاویر پاتولوژیک عملکرد قابل توجهی را نشان دهد. این روش عملکرد بهتری در مقایسه ارزیابی توسط معیار میانگین متوسط دقت نسبت به سایر روش‌ها داشته است و همچنین میدانیم که روش توضیح داده شده در بالا [۱۲] تصویر ورودی را به شبکه عصبی بسیار سریع مبتنی بر منطقه^۲ می‌دهد و شبکه تصویر را به لایه های کانولوشنالی تبدیل کرده و سپس عمل کلاسه بندی و طبقه بندی را انجام می‌دهد. آزمایش روی مجموعه داده تصویر پزشکی اسکن ریه نشان می‌دهد که روش ما می‌تواند ابهام مناطق غیرطبیعی پاتولوژیک را کاهش دهد و روش پیشنهادی این پتانسیل را دارد که در سیستم‌های تشخیصی پیشرفته پزشکی به کار گرفته شود و تحقیقات آتی می‌تواند به بهبود بیشتر دقت و سرعت این مدل در مواجهه با داده‌های پزشکی چالش‌برانگیز کمک کند.

¹ Attention Classifier

² Fast R-CNN

مراجع

- [1] Tang, Hai, and Zhihui Hu. "Research on medical image classification based on machine learning." IEEE access 8 (2020)
- [2] Agrawal, Shubham, et al. "Content-based medical image retrieval system for lung diseases using deep CNNs." International Journal of Information Technology (2022)
- [3] Fang, Jiansheng, Huazhu Fu, and Jiang Liu. "Deep triplet hashing network for case-based medical image retrieval." Medical image analysis 69 (2021)
- [4] J. Kim and S.-E. Yoon, "Regional attention based deep feature for image retrieval.," in BMVC, p. 209, 2018.
- [5] Ahmed Ahmed, Ali, and Sharaf J. Malebary. "Query expansion based on top-ranked images for content-based medical image retrieval." IEEE Access 8 (2020)
- [6] Ahmed Ahmed, Ali. "Implementing relevance feedback for content-based medical image retrieval." IEEE Access 8 (2020)
- [7] T. Ng, V. Balntas, Y. Tian, and K. Mikolajczyk, "Solar: Second-order loss and attention for image retrieval," arXiv preprint arXiv:2001.08972, 2020.
- [8] E. Yang, D. Yao, B. Cao, H. Guan, P.-T. Yap, D. Shen, and M. Liu, "Deep disentangled hashing with momentum triplets for neuroimage search," in International Conference on Medical Image Computing and Computer Assisted Intervention, pp. 191–201, Springer, 2020.
- [9] Fang, Jiansheng, et al. "Combating ambiguity for hash-code learning in medical instance retrieval." IEEE Journal of Biomedical and Health Informatics (2021)
- [10] G. Toliás, R. Sivic, and H. Jegou, "Particular object retrieval with integral ℓ^2 max-pooling of cnn activations," arXiv preprint arXiv:1511.05879, 2015.
- [11] Xie, Lingxi, Tian, Q, Hong, R, and Zhang, B. Image classification and retrieval are one. In ICMR, 2015.
- [12] Anwar, Mohd Ayaan, et al. "Attention to Traffic: Network Traffic Classification using Attention-Based CNNs." 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT). IEEE, 2023.
- [13] Liu, Yunhe, et al. "Attention-based deep multiple-instance learning for classifying circular RNA and other Long non-coding RNA." *Genes* 12.12 (2021)
- [14] Zhao, Gangming, Weifeng Ge, and Yizhou Yu. "GraphFPN: Graph feature pyramid network for object detection." Proceedings of the IEEE/CVF international conference on computer vision. 2021.
- [15] Malhotra, Priyanka, et al. "Thoracic Disease Chest Radiographic Image Dataset: A Comprehensive Review." International Conference on Intelligent Systems Design and Applications. Cham: Springer International Publishing, 2021.