

بهبود قطعه بندی تصاویر پزشکی با استفاده از یادگیری ماشین: نقش بهینه سازی ویژگی های استخراج شده در شبکه های عصبی کانولوشنی

مسلم کاویانی^۱ | ام کلثوم شهریاری^{۲*}

چکیده:

در سال های اخیر، شبکه های عصبی کانولوشنی مبتنی بر یادگیری عمیق به پیشرفت های چشمگیری، به ویژه در حوزه قطعه بندی تصاویر پزشکی، دست یافته اند. با این حال، عواملی مانند طراحی یکنواخت لایه ها، بهره گیری ناکافی از اطلاعات چندمقیاسی و افزایش پیچیدگی مدل ها به دلیل تعداد بالای پارامترها، عملکرد این روش ها را در شرایط پیچیده محدود کرده است. افزون بر این، امنیت داده ها و حفظ حریم خصوصی از چالش های اساسی در پردازش داده های پزشکی محسوب می شوند.

در این پژوهش، یک مدل کانولوشنی رمزگذار-رمزگشا بهبود یافته معرفی شده است که از استراتژی هایی برای بهینه سازی ویژگی های استخراج شده و کاهش تعداد پارامترها بهره می برد. این مدل با به کارگیری سازوکارهای حفظ اطلاعات پایه و ماژول های توجه متراکم، توانایی استخراج اطلاعات چندسطحی را ارتقا داده و به عنوان یک گزینه بهینه برای استفاده در ساختار یادگیری فدرال، امنیت و کارایی بیشتری را ارائه می دهد. ارزیابی مدل پیشنهادی بر روی مجموعه داده ClinicDB-CVC نشان می دهد که این روش در مقایسه با سایر روش های پیشرفته، عملکرد بهتری را از نظر معیار میانگین تقاطع بر روی اتحاد ارائه می دهد.

کلید واژه ها: یادگیری ماشین، قطعه بندی تصاویر پزشکی، یادگیری فدرال، معماری قطعه بندی عمیق با فشردگی ویژگی ها

^۱گروه مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه آزاد اسلامی واحد سنجند، سنجند، ایران، mo.kaviani@iau.ac.ir

^۲گروه مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه آزاد اسلامی واحد سنجند، سنجند، ایران، shahryari.k@iausdj.ac.ir

نویسنده مسئول

ام کلثوم شهریاری، استادیار گروه مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه آزاد اسلامی واحد سنجند، سنجند، ایران، shahryari.k@iausdj.ac.ir

تاریخ دریافت: ۱۴۰۳/۱۲/۲۵

تاریخ پذیرش: ۱۴۰۴/۳/۷

۱-مقدمه

تحلیل تصاویر پزشکی نقش مهمی در تشخیص و درمان سرطان ایفا می کند. تصاویر میکروسکوپی می توانند اطلاعات ارزشمندی درباره مراحل بیماری، تمایز انواع تومورها، مکانیسم های ژنتیکی سلولی و مولکولی، و همچنین کاربردهای مرتبط با بیماری های مزمن و سرطان ارائه دهند [۱]. به طور سنتی، تحلیل این تصاویر به صورت دستی توسط پاتولوژیست ها انجام می شود، فرآیندی که وابسته به تجربه پزشکان است، زمان زیادی می برد و در عین حال مستعد خطا و وابستگی به فرد است. از این رو، سیستم های

تشخیص رایانه‌ای به دلیل نقش کلیدی خود در تشخیص پزشکی و تحلیل تصاویر، توجه بسیاری از محققان و کاربردهای بالینی را به خود جلب کرده‌اند. این سیستم‌ها عمدتاً بر اساس قطعه‌بندی تصاویر پزشکی عمل می‌کنند [۲]. برخلاف وظایف طبقه‌بندی و تشخیص، هدف اصلی قطعه‌بندی تصاویر پزشکی، جداسازی دقیق اشیای خاص از پس‌زمینه تصویر است که منجر به تحلیل دقیق‌تر بیماری‌ها می‌شود [۳].

الگوریتم‌های کلاسیک و فازی مورد استفاده در قطعه‌بندی تصاویر پزشکی هر یک دارای پارامترهای خاصی هستند که آن‌ها را برای کاربردهای متفاوت مناسب می‌سازد. با این حال، این روش‌ها معمولاً در تعمیم‌پذیری روی داده‌های پیچیده با محدودیت‌هایی مواجه‌اند. علاوه بر این، کیفیت تصاویر پزشکی تأثیر مستقیمی بر عملکرد این الگوریتم‌ها دارد. بنابراین، توسعه مدل‌هایی که در برابر چالش‌های مختلف مقاومت بالایی داشته باشند، امری ضروری است [۴].

در سال‌های اخیر، شبکه‌های عصبی کانولوشنی با معماری‌های رمزگذار-رمزگشا توانسته‌اند عملکرد بهتری در قطعه‌بندی تصاویر پزشکی نسبت به روش‌های سنتی ارائه دهند [۵]. موفقیت این مدل‌ها عمدتاً به دلیل بهره‌گیری از اتصال‌های میان‌بر است که اطلاعات معنایی سطح پایین را با اطلاعات سطح بالا ترکیب کرده و به تولید ماسک نهایی کمک می‌کند [۶]. با این حال، بسیاری از معماری‌های پیشرفته تنها بر بهینه‌سازی الگوریتم‌های استخراج ویژگی‌های عمیق متمرکز هستند و اهمیت اطلاعات با وضوح بالا در مراحل اولیه رمزگذاری را نادیده می‌گیرند. این اطلاعات می‌توانند با جبران افت اطلاعات مکانی که طی عملیات کاهش اندازه رخ می‌دهد، دقت مدل را بهبود بخشند [۷].

در این مقاله، یک معماری رمزگذار-رمزگشا برای قطعه‌بندی تصاویر پزشکی با عنوان شبکه قطعه‌بندی عمیق با فشردگی ویژگی‌ها معرفی شده است. این معماری با هدف توسعه مدلی سبک و فشرده طراحی شده است که ضمن حفظ کارایی، تعداد پارامترها و پیچیدگی را کاهش دهد.

در بخش رمزگذار، مدل پیشنهادی از مکانیزم حفاظت اطلاعات پایه استفاده می‌کند که علاوه بر کاهش تعداد پارامترها، اطلاعات مکانی بلندمدت را در لایه‌های معنایی سطح پایین حفظ کرده و به بلوک توجه تقسیم فشرده ارسال می‌کند. این بلوک با بهره‌گیری از یک ساختار توجه چندمسیره، نمایش ویژگی‌های کانال‌های مختلف را تقویت می‌کند.

در بخش رمزگشا، ویژگی‌های رمزگذاری شده در هر لایه کاهش‌یافته با ویژگی‌های افزایش‌یافته از طریق اتصال‌های میان‌بر ادغام می‌شوند و همان بلوک توجه تقسیم فشرده برای استخراج ویژگی‌های ترکیبی به کار گرفته می‌شود.

طراحی این مدل به گونه‌ای است که در مواجهه با داده‌های پیچیده، عملکرد پایداری داشته باشد. همچنین، کاهش وابستگی به کیفیت تصاویر ورودی، آن را برای کاربردهای پزشکی مختلف، از جمله داده‌های پرنویز یا با کنتراست پایین، قابل اعتماد می‌سازد.

ارزیابی مدل پیشنهادی روی مجموعه داده ClinicDB-CVC [۸] انجام شده است. نتایج این ارزیابی نشان می‌دهد که مدل پیشنهادی در مقایسه با سایر روش‌های پیشرفته، عملکرد بهتری از نظر معیارهای میانگین تقاطع بر روی اتحاد و ضریب دایس ارائه می‌دهد.

۱-۱- انگیزه پژوهش

با پیشرفت فناوری و افزایش استفاده از سیستم‌های تحلیل تصاویر پزشکی، نیاز به روش‌های دقیق و کارآمد برای قطعه‌بندی این تصاویر بیش از گذشته احساس می‌شود. روش‌های سنتی مانند آستانه‌گذاری و الگوریتم‌های مبتنی بر ویژگی‌های پیکسلی،

در مواجهه با تصاویر پیچیده و کم کیفیت دچار محدودیت هایی هستند که می تواند به تشخیص های نادرست و در نهایت درمان های ناموفق منجر شود.

علاوه بر این، با رشد فزاینده حجم داده های پزشکی، تبادل این اطلاعات میان مراکز درمانی و پژوهشی برای بهبود مدل های یادگیری ماشینی امری ضروری است. با این حال، چنین تبدیلی چالش های قابل توجهی را در زمینه حفظ امنیت و حریم خصوصی داده ها به همراه دارد. داده های پزشکی به دلیل حساسیت بالای خود، در صورت افشا می توانند پیامدهای جدی قانونی و اخلاقی داشته باشند. از این رو، توسعه روش هایی که علاوه بر بهبود دقت قطعه بندی، امنیت داده ها را نیز تضمین کنند، یک ضرورت محسوب می شود.

این پژوهش با هدف رفع این خلأ انجام شده است و تلاش می کند مدلی پیشرفته و ایمن ارائه دهد که ضمن حفظ حریم خصوصی داده ها، عملکرد بهینه ای در قطعه بندی تصاویر پزشکی داشته باشد. این مدل باید قادر باشد با چالش های مرتبط با داده های پیچیده، متنوع و پرنویز سازگار بوده و در کاربردهای گوناگون پزشکی مورد استفاده قرار گیرد.

۱-۲-هدف پژوهش

هدف اصلی این پژوهش، طراحی و ارائه یک مدل پیشرفته مبتنی بر معماری رمزگذار-رمزگشا است که با بهره گیری از حفظ دقت قطعه بندی تصاویر پزشکی، از امنیت داده ها و حریم خصوصی بیماران نیز محافظت کند. مدل پیشنهادی با استفاده از تکنیک های نوینی مانند مکانیزم حفاظت اطلاعات پایه و بلوک توجه تقسیم فشرده، توانایی مقابله با چالش های تصاویر پیچیده را افزایش داده و عملکرد مطلوبی را در سناریوهای مختلف پزشکی ارائه می دهد.

۲-روش تحقیق

روش های یادگیری عمیق مبتنی بر شبکه های عصبی کانولوشنی توانسته اند عملکرد قابل توجهی در قطعه بندی تصاویر پزشکی ارائه دهند. معماری شبکه های کانولوشنی رمزگذار-رمزگشا [۹] که شامل دو بخش رمزگذار و رمزگشا است، در قطعه بندی انواع تصاویر پزشکی نتایج مطلوبی به همراه داشته است. کائور و همکاران [۱۰] نسخه ای پیشرفته از این مدل تحت عنوان U-Net++ را معرفی کردند که با اضافه کردن مسیرهای تو در تو و اتصال های میان بر، میزان از دست رفتن اطلاعات معنایی را در هنگام ترکیب ویژگی های رمزگذار و رمزگشا کاهش می دهد. در همین راستا، ژو و همکاران [۱۱] رویکردی مبتنی بر اتصالات میان بر تمام مقیاس ارائه دادند که اطلاعات با وضوح بالا و پایین را در سطوح مختلف ترکیب می کند.

DoubleU-Net، که توسط ژا و همکاران [۱۲] توسعه یافت، از دو شبکه کانولوشنی رمزگذار-رمزگشا تشکیل شده است که به صورت ترتیبی عمل می کنند. در این مدل، بخش رمزگذار وظیفه استخراج و تجزیه و تحلیل ویژگی های پیچیده تصویر را بر عهده دارد و در پایان هر مرحله از کاهش نمونه برداری، اطلاعات متنی را جمع آوری می کند.

برای حل مشکل محو شدن گرادیان در شبکه های عمیق، شن و همکاران [۱۳] معماری ResNet را معرفی کردند که در بسیاری از مدل های قطعه بندی مورد استفاده قرار گرفته است. در زمینه قطعه بندی پزشکی، ژا و همکاران [۱۴] نسخه پیشرفته ای از این معماری به نام ResUNet++ را برای قطعه بندی پولیپ ارائه دادند که شامل ماژول فشرده سازی و تحریک و مکانیزم توجه است. دازا و همکاران [۱۵] نیز نسخه های سبک وزنی از معماری شبکه های کانولوشنی رمزگذار-رمزگشا را برای قطعه بندی تومور مغزی طراحی کردند که در آن، از ماژول های گسترش و بلوک های متراکم در بخش رمزگذار و رمزگشا استفاده شده است.

برخلاف تلاش‌هایی که بر بهینه‌سازی معماری شبکه متمرکز بودند، سابروسکی و همکاران [۱۶] مدل nnU-Net را معرفی کردند که به جای تغییر در ساختار شبکه، اهمیت پیش‌پردازش داده‌ها، استراتژی‌های آموزش مدل و فرایند استنتاج نتایج را برجسته می‌کند.

۲-۱- مازول‌های توجه متراکم

در سال‌های اخیر، مکانیزم توجه به‌سرعت در حوزه بینایی کامپیوتر گسترش یافته است و به‌عنوان یکی از رویکردهای کلیدی در قطعه‌بندی تصاویر پزشکی مورد استفاده قرار گرفته است. یکی از این مدل‌ها، مکانیزم توجه کانالی است که شامل یک مازول فشرده‌سازی برای جمع‌آوری اطلاعات مکانی کلی و یک مازول تحریک برای استخراج روابط بین کانال‌های نقشه ویژگی می‌شود [۱۷].

علاوه بر توجه کانالی، توجه مکانی نیز به‌عنوان یک مکانیزم انتخاب تطبیقی موقعیت‌های مکانی شناخته می‌شود. به‌عنوان مثال، لو چنگ [۱۸] مدل شبکه‌های رمزگذار-رمزگشای توجهی را معرفی کرد که از یک دروازه توجه پایین به بالا برای تمرکز دقیق‌تر بر نواحی خاص تصویر بهره می‌برد.

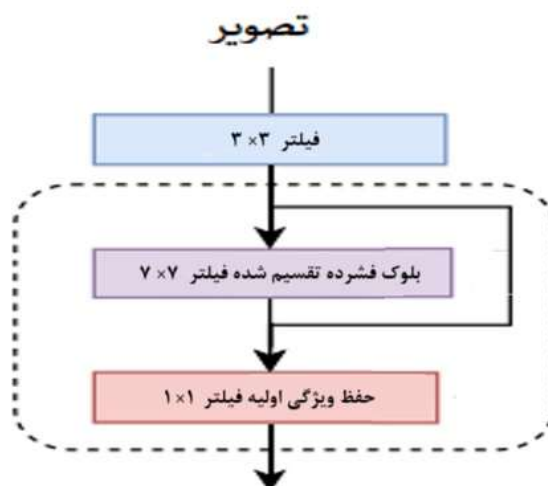
علاوه بر این، ادغام مدل‌های انتقالی با شبکه‌های عصبی کانولوشنی عملکرد قابل توجهی در قطعه‌بندی تصاویر پزشکی نشان داده است. برای نمونه، باسیال و همکاران [۱۹] مدل شبکه عصبی انتقالی پیچشی را پیشنهاد کردند که در آن، شبکه‌های عصبی کانولوشنی برای استخراج پیچ‌های تصویری در نیمه اول رمزگذار به کار گرفته شده و مدل‌های انتقالی برای استخراج زمینه‌های جهانی استفاده می‌شوند.

۲-۲- مازول دورانی پیشرفته

ماژول دورانی پیشرفته یکی از معماری‌های کارآمد در شبکه‌های عصبی است که توسط لی و همکاران [۲۰] معرفی شد. در این ساختار، هر فیلتر کانولوشن مسئول پردازش یک کانال ورودی است. برخلاف کانولوشن استاندارد، این روش قادر است عملکرد مشابهی را با کاهش قابل توجه تعداد پارامترها و محاسبات ارائه دهد.

مدل MobileNets با بهره‌گیری از این تکنیک، به‌عنوان یک معماری نوآورانه در یادگیری عمیق مطرح شد. در زمینه دسته‌بندی تصاویر، پنگیو و همکاران [۲۱] از مازول دورانی پیشرفته برای توسعه مدل‌هایی استفاده کردند که نه تنها از روش‌های پیشین کارآمدتر بودند، بلکه پیچیدگی محاسباتی کمتری نیز داشتند.

با این حال، واتس و همکاران [۲۲] دریافتند که مازول دورانی پیشرفته در نقشه‌های ویژگی با تعداد کانال‌های کم عملکرد مطلوبی ندارد. برای رفع این محدودیت، آن‌ها مدل MobileNetV2 را ارائه دادند که در آن، یک کانولوشن 1×1 پیش از عملیات دورانی پیشرفته به کار گرفته می‌شود تا ابعاد ویژگی‌ها افزایش یابد و عملکرد مدل بهبود پیدا کند.



شکل ۱: مدل حفظ اطلاعات پایه

۳-مدل پیشنهادی

۳-۱-حفظ اطلاعات پایه

در اکثر مدل های قطعه بندی تصاویر پزشکی، عملیات کانولوشن در اولین بلوک نمونه برداری به پایین برای استخراج اطلاعات معنایی سطح پایین از تصاویر استفاده می شود. معماری مدل کانولوشنی رمزگذار-رمزگشا به طور گسترده در مدل های مختلف مورد استفاده قرار گرفته است. مدل بلاک اولیه معمولاً به گونه ای طراحی می شود که میدان دید مؤثر مشابه با کانولوشن 7×7 را فراهم کند و تعداد پارامترها را کاهش دهد.

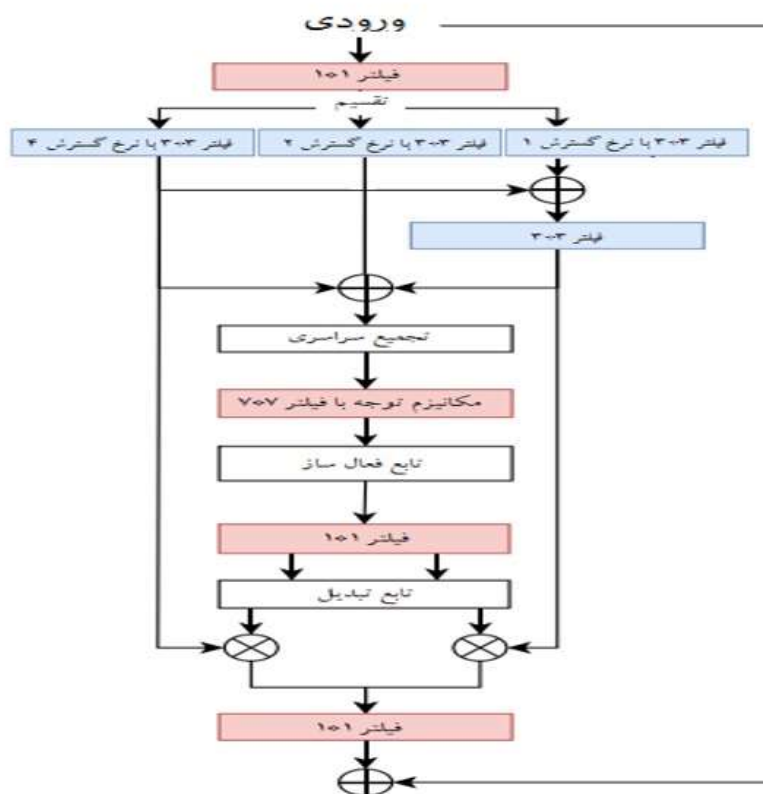
در مدل ResUNet++ لایه اول نمونه برداری به پایین شامل استراتژی اتصال میان بری است تا اثرات منفی ناپدید شدن گرادیان را کاهش دهد، هرچند افزودن بلوک های کانولوشنی بیشتر می تواند میدان دید مؤثر شبکه را گسترش دهد، اما این کار منجر به افزایش سریع تعداد پارامترها و محاسبات می شود و ممکن است پایداری مدل را تحت تأثیر قرار دهد. همچنین، تحقیقات اخیر نشان داده اند که میدان دید مؤثر در صورت افزایش تعداد کانولوشن های 3×3 پیایی، ممکن است تا حدی کاهش یابد.

برای رفع این مسئله، ما یک استراتژی جدید به نام حفظ اطلاعات پایه را در اولین بلوک نمونه برداری به پایین معرفی می کنیم که در شکل ۱ ارائه شده است. اصلاح اصلی این ماژول استفاده از کانولوشن جداسازی عمقی است، که شامل یک کانولوشن عمقی 7×7 به دنبال یک کانولوشن نقطه ای 1×1 می باشد. از آنجا که کانولوشن جداسازی عمقی هزینه محاسباتی و تعداد پارامترها را در مقایسه با کانولوشن استاندارد کاهش می دهد، این امکان فراهم می شود که از اندازه های بزرگ هسته در کانولوشن عمقی استفاده کنیم تا اطلاعات مکانی دوردست را ادغام کرده و اطلاعات پایه را تا حد امکان در لایه معنایی سطح پایین حفظ کنیم. برای بهبود ویژگی های کانال، یک کانولوشن 3×3 در ابتدای این ماژول اضافه شده است، زیرا عملکرد کانولوشن جداسازی عمقی در ویژگی های کم بعد کاهش می یابد در نهایت، ماژول پیشنهادی حفظ اطلاعات پایه می تواند بدون افزایش تعداد پارامترها و هزینه محاسباتی، عملکرد شبکه را بهبود بخشد.

۳-۲- مازول فشرده سازی توزیع شده

مدل هایی مانند DoubleUnet و ساختارهای مبتنی بر اتصالات باقیمانده در بسیاری از روش های قطعه بندی معنایی مورد استفاده قرار گرفته اند. با این حال، در برخی از شبکه های کلاسیک، لایه های کانولوشنی به صورت متوالی روی هم قرار می گیرند که این موضوع باعث می شود هر لایه تنها میدان دید ثابتی داشته باشد. در قطعه بندی تصاویر پزشکی، ضایعات می توانند اندازه های متنوعی داشته باشند، بنابراین نمایش ویژگی های چندمقیاسی نقش مهمی در بهبود عملکرد مدل ایفا می کند. به همین دلیل، در سال های اخیر رویکردهای جدیدی برای بهره گیری از ویژگی های بین کانالی در فرآیند یادگیری مدل ها پیشنهاد شده اند [۲۳].

با الهام از این روش ها، ما معماری جدیدی به نام مازول فشرده سازی توزیع شده توسعه داده ایم. یک نمای کلی از بلوک پیشنهادی جدید در شکل ۲ نشان داده شده است. این مدل برای بهینه سازی استخراج ویژگی های چندمقیاسی و کاهش تعداد پارامترها، از سه گروه با نرخ گسترش متفاوت $d=1$ ، $d=2$ و $d=4$ استفاده می کند. این سه گروه از ویژگی های ورودی تقسیم شده، ابتدا از یک فیلتر کانولوشنی 1×1 برای کاهش ابعاد عبور می کنند و سپس وارد فیلترهای کانولوشن های 3×3 با نرخ گسترش های متفاوت می شوند تا ویژگی های چندمقیاسی را استخراج کنند. ویژگی های خروجی هر سه گروه با یکدیگر جمع شده و وارد یک مکانیسم توجه فضایی با کانولوشن 7×7 می شوند که اطلاعات مناطق مهم نقشه ویژگی را تقویت می کند. همچنین، یک اتصال باقی مانده همراه با تطابق ابعاد با استفاده از کانولوشن 1×1 به خروجی نهایی اضافه می شود. این طراحی جدید توانایی بالاتری در استخراج اطلاعات جهانی و محلی از نقشه های ویژگی ارائه می دهد و با پردازش چندمقیاسی، عملکرد بهتری در قطعه بندی تصاویر پزشکی به ویژه برای شناسایی ضایعات با اندازه های مختلف دارد.



شکل ۲: مازول فشرده سازی توزیع شده

۳-۳- معماری مدل پیشنهادی

در این پژوهش، مدلی مبتنی بر استراتژی حفظ اطلاعات پایه و بلوک فشرده تقسیم شده ارائه شده است که از معماری رمزگذار- رمزگشا پیروی کرده و تحت عنوان شبکه قطعه بندی عمیق با فشرده سازی ویژگی ها شناخته می شود. طرح کلی این مدل در شکل ۳ نمایش داده شده است.

بخش رمزگذار: در مرحله رمزگذاری، استراتژی حفظ اطلاعات پایه برای استخراج اطلاعات معنایی سطح پایین از تصاویر ورودی مورد استفاده قرار می گیرد. در این فرآیند، کانولوشن عمقی جداسازی با یک کرنل 7×7 به کار گرفته می شود تا میدان دید شبکه گسترش یابد و اطلاعات پایه بدون افزایش تعداد پارامترها حفظ شوند.

ماژول فشرده سازی توزیع شده، متشکل از گروه های ویژگی چندمسیره با تعداد متنوعی از کانولوشن ها و مکانیزم توجه است که اطلاعات کانالی را در مقیاس های مختلف میدان دید ترکیب کرده و ویژگی های معنایی معنادار را برجسته می سازد. در ادامه، هر بلوک با عملیات ماکس پولینگ 2×2 همراه با گام ۲ دنبال می شود تا عملیات کاهش مقیاس صورت گیرد.

بخش رمزگشا: در مرحله رمزگشایی، هر زیرشبکه ابتدا با یک عملگر افزایش مقیاس آغاز می شود که به تدریج اندازه تصویر ورودی را بازیابی می کند. برای حفظ اطلاعات کلیدی، اتصالات میان بر میان نقشه های ویژگی رمزگشا و لایه های متناظر رمزگذار برقرار می شود. این فرایند باعث می شود که اطلاعات سطح پایین و سطح بالا ترکیب شوند و در نتیجه، ماسک خروجی با دقت بالاتری ایجاد شود.

پس از این مرحله، بلوک های فشرده تقسیم شده اضافه می شوند تا مشکل ناپدید شدن گرادین برطرف شده و ویژگی های مؤثرتری استخراج شوند. در مرحله نهایی، یک کانولوشن 1×1 به همراه لایه سیگموئید یا سافت ماکس به کار گرفته می شود تا ماسک قطعه بندی باینری یا چندکلاسه تولید گردد.

این طراحی باعث می شود که مدل یادگیری عمیق برای قطعه بندی با ادغام ویژگی ها از توانایی بالایی در استخراج و ترکیب اطلاعات معنایی تصاویر پزشکی برخوردار باشد و نتایج دقیقی ارائه دهد. علاوه بر این، ساختار این مدل به گونه ای بهینه سازی شده است که برای یادگیری فدرال مناسب باشد. با بهره گیری از مدل داده های فشرده و کاهش تعداد پارامترها، زیرساختی کارآمد برای پردازش غیرمتمرکز داده های پزشکی فراهم می شود.



۴-۱-دیتاست

۴-۲- معیارهای ارزیابی

برای ارزیابی عملکرد مدل‌ها در قطعه‌بندی تصاویر پزشکی، از معیارهای استاندارد مانند میانگین تقاطع بر اتحاد (mIoU)، دقت (Accuracy)، یادآوری (Recall) و دقت پیش‌بینی (Precision) استفاده می‌شود. میانگین تقاطع بر اتحاد (mIoU) به‌عنوان یک معیار رایج در مسابقات برای مقایسه عملکرد مدل‌ها مورد استفاده قرار می‌گیرد.

نتایج کمی روی دیتاست CVC-ClinicDB در جدول ۱ نمایش داده شده است. در قطعه بندی تصاویر پزشکی، عملکرد شبکه بر اساس متریک mIoU معمولاً بیشتر مورد توجه قرار می گیرد. از جدول ۱ مشخص می شود که مدل حفظ اطلاعات پایه با بلوک فشرده موفق به بهبود عملکرد قابل توجهی در مقایسه با مدل های مشابه شده است.

جدول ۱: مقایسه مدل پیشنهادی با چند مدل مطرح بر روی دیتاست Cvc-Clinic-DB

ردیف	مدل ها	Accuracy	Precision	Recall	mIoU
۱	R2 U-Net	۰.۹۷۸	۰.۸۸۰	۰.۸۴۷	۰.۷۶۵
۲	Levit U-Net	۰.۹۸۰	۰.۸۴۹	۰.۸۲۶	۰.۷۵۴
۳	Trans U-Net	۰.۹۸۲	۰.۸۷۶	۰.۸۷۳	۰.۷۹۹
۴	Res U-Net++	۰.۹۸۲	۰.۸۷۰	۰.۸۵۳	۰.۷۸۱
۵	U-Net3	۰.۹۸۴	۰.۹۰۷	۰.۸۸۵	۰.۸۲۷
۶	U-Net++	۰.۹۸۴	۰.۹۱۹	۰.۸۵۹	۰.۸۱۱
۷	U-Net	۰.۹۸۴	۰.۸۸۲	۰.۸۹۳	۰.۸۰۹
۸	Att U-Net	۰.۹۸۶	۰.۹۰۴	۰.۹۰۱	۰.۸۳۵
۹	Double U-Net	۰.۹۸۶	۰.۸۹۲	۰.۹۱۲	۰.۸۳۶
۱۰	مدل پیشنهادی	۰.۹۹۰	۰.۹۱۷	۰.۹۳۰	۰.۸۶۱

در یادگیری فدرال و با تأکید بر حفظ امنیت داده ها، استفاده از معماری های رمزگذار-رمزگشا برای قطعه بندی معنایی تصاویر پزشکی به دلیل ویژگی های قدرتمند در استخراج و تحلیل اطلاعات، امری رایج است. روش پیشنهادی ما، با استفاده از کانولوشن های عمقی جادشده با کرنل های بزرگ، اطلاعات پایه تصاویر را حفظ کرده و بر اساس مازول فشرده سازی توزیع شده و مکانیزم حفظ اطلاعات پایه برای کاهش پهنای باند و هزینه های ارتباطی، وزن ها و گرادین های ارسالی از کلاینت ها به سرور به صورت فشرده کاربرد دارد. امکان ارسال سبک و امن پارامترها را فراهم می کنند. این فشرده سازی بدون افت دقت مدل، کارایی ارتباطات را بهبود می بخشد.

ویژگی حریم خصوصی به ویژه در سناریوهای پزشکی که حریم خصوصی بیماران اهمیت بالایی دارد، کارایی خود را نشان می دهد. استفاده از این روش ها در یادگیری فدرال، تضمین کننده امنیت داده ها بدون کاهش عملکرد مدل است. در فرآیند یادگیری فدرال، امنیت ارتباطات بین کلاینت ها و سرور مرکزی از اهمیت بالایی برخوردار است. رمزنگاری انتها به انتها اطمینان می دهد که داده ها و مدل ها در طول انتقال غیرقابل دسترسی باشند. علاوه بر این، تکنیک های جمع آوری امن امکان ترکیب مدل های محلی را بدون افشای اطلاعات هر کلاینت فراهم می کنند. این روش ها، حریم خصوصی و امنیت داده ها را به طور همزمان تضمین می کنند.

نوآوری های اصلی این تحقیق، طراحی مازول فشرده سازی توزیع شده است که با تقویت اتصال میان کانال ها و به کارگیری مکانیزم توجه، ویژگی های چندمقیاسی را بهینه می کند. این بلوک ها در محیط یادگیری فدرال بسیار مؤثر هستند، زیرا می توانند اطلاعات عمیق را بدون انتقال داده های حساس بین کلاینت ها و سرور استخراج کنند.

مدل پیشنهادی ما، با تعداد پارامترهای کمتر و زمان استنتاج قابل قبول، برای پیاده سازی در دستگاه های با حافظه محدود و در شرایط یادگیری فدرال مناسب است. همچنین، این مدل می تواند در سناریوهایی که تصاویر دارای کیفیت پایین، نویز بالا یا شرایط چالش برانگیز هستند، عملکرد قابل توجهی داشته باشد.

برای آینده، چندین مسیر پژوهشی پیشنهاد می شود:

-گسترش چارچوب به تصاویر سه بعدی پزشکی برای بهره برداری

-ارزیابی عملکرد یادگیری فدرال روی داده های واقعی بیمارستانی، به منظور تحلیل اثربخشی در سناریوهای واقعی.

-استفاده از الگوریتم های ترکیب پیشرفته برای مدیریت داده های غیرهمگن بین کلاینت ها.

-ادغام یادگیری فدرال با مدل های انتقالی برای بهبود دقت در سناریوهای با داده های کم.

به طور کلی، مدل پیشنهادی ما نشان دهنده عملکرد برجسته ای در قطعه بندی تصاویر پزشکی است و می تواند به عنوان یک مدل امن و کارآمد برای کاربردهای مختلف در محیط های یادگیری فدرال مورد استفاده قرار گیرد.

مراجع

- [1] Abdel-Nabi. Heba, Ali. Mostafa and Awajan. Arafat, "A comprehensive review of the deep learning-based tumor analysis approaches in histopathological images: segmentation, classification and multi-learning tasks," Cluster Computing Volume 26, Issue 5, Pages 3145 - 3185 October 2023 , doi: 10.1007/s10586-022-03951-2
- [2] Manju A. Arivukarasi, "AEDAMIDL: An Enhanced and Discriminant Analysis of Medical Images using Deep Learning," Proceedings of the 3rd International Conference on Smart Technologies in Computing, Electrical and Electronics, ICSTCEE 2022 Bengaluru 16 December 2022 through 17 December 2022 , doi: 10.1109/ICSTCEE56972.2022.10100240
- [3] Yang. Mengzhu, Wang. Yongfang and Li, Guoqiang, "SU-Net: A retinal segmentation model based on improved U-Net network," ACM International Conference Proceeding Series Pages 946 - 950 16 December 2022, doi: 10.1145/3584376.3584545
- [4] Shu. Xiu, Yang. Yunyun and Liu. Jun, "ALVLS: Adaptive local variances-Based levelset framework for medical images segmentation," Pattern Recognition Volume 136 April 2023, doi: 10.1016/j.patcog.2022.109257
- [5] Hussain. Tahir and Shouno. Hayaru, "MAGRes-UNet: Improved Medical Image Segmentation Through a Deep Learning Paradigm of Multi-Attention Gated Residual U-Net," IEEE Access Open Access Volume 12, Pages 40290 – 40310 2024, doi: 10.1109/ACCESS.2024.3374108
- [6] Choubineh. Abouzar, Chen. Jie and Coenen. Frans, "A Quantitative Insight Into the Role of Skip Connections in Deep Neural Networks of Low Complexity: A Case Study Directed at Fluid Flow Modeling," Journal of Computing and Information Science in Engineering Volume 23, Issue 1 February 2023, doi: 10.1115/1.4054868
- [7] Saoudi. Rania, Boudechiche. Djameleddine and Messali. Zoubeida, "Brain MRI Scans Super-Resolution With Wavelet and Attention Mechanisms," 2nd International Conference on Electrical Engineering and Automatic Control, ICEEAC 2024, doi: 10.1109/ICEEAC61226.2024.10576395
- [8] Caicedo. Juan et al, Cimini, Beth A, "Nucleus segmentation across imaging experiments: the 2018 Data Science Bowl," Nature Methods Open Access Volume 2019, doi: 10.1038/s41592-019-0612-7
- [9] Zhang. Rong, Zhang. Rongguo and Ma. Jiechao, "Analysis of different encoder-decoder-based approaches for biomedical imaging segmentation," ACM International Conference Proceeding Series Pages 105 - 113 20 November 2020 Article number 3449320 6th International Conference on Robotics and Artificial Intelligence, ICRAI 2020, doi: 10.1145/3449301.3449320

- [10] Kaur Buttar. Preetpal and Sachan. Manoj Kumar, "Semantic segmentation of clouds in satellite images based on U-Net++ architecture and attention mechanism," Expert Systems with ApplicationsVolume 20915 December 2022, doi: 10.1016/j.eswa.2022.118380
- [11] Xu. Hanwen, Tang. Xinming and Yang. Fanlin, "Feature-Selection High-Resolution Network With Hypersphere Embedding for Semantic Segmentation of VHR Remote Sensing Images," IEEE Transactions on Geoscience and Remote SensingVolume 602022, doi: 10.1109/TGRS.2022.3183144
- [12] Jha. Debesh and Riegler. Michael A, "DoubleU-Net: A deep convolutional neural network for medical image segmentation," Proceedings - IEEE Symposium on Computer-Based Medical SystemsOpen AccessVolume 2020-July, Pages 558 – 564, doi: 10.1109/CBMS49503.2020.00111
- [13] Shen. Zhixi and Liu. Yong, "A novel connectivity of deep convolutional neural networks," Proceedings - 2017 Chinese Automation Congress, CAC 2017Volume 2017-January, Pages 7779 - 778329 December 2017, doi: 10.1109/CAC.2017.8244187
- [14] Jha. Debesh et al, "A Comprehensive Study on Colorectal Polyp Segmentation with ResUNet++, Conditional Random Field and Test-Time Augmentation," IEEE Journal of Biomedical and Health InformaticsOpen AccessVolume 25, Issue 6, Pages 2029 - 2040June 2021, doi: 10.1109/JBHI.2021.3049304
- [15] Daza. Laura and Gómez. Catalina, "Cerberus: A Multi-headed Network for Brain Tumor Segmentation," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)Volume 12659 LNCS, Pages 342 – 3512021, doi: 10.1007/978-3-030-72087-2_30
- [16] Sabrowsky. Hirsch, Bertram. Thumfart, "A Content-Driven Architecture for Medical Image Segmentation," ACM International Conference Proceeding SeriesPages 89 - 9627 November 2020, doi: 10.1145/3442555.3442570
- [17] Wang. Zekun, Liu. Peter X and Zou, Yanni, "Hybrid dilation and attention residual U-Net for medical image segmentation," Computers in Biology and MedicineVolume 134July 2021, doi: 10.1016/j.compbiomed.2021.104449
- [18] Lu. Chengsong, "Performance analysis of attention mechanism and teacher forcing ratio based on machine translation," Journal of Physics: Conference SeriesOpen AccessVolume 2580, Issue 12023, doi: 10.1088/1742-6596/2580/1/012006
- [19] Basyal. Ganga et al, "Comparative study of CNN models for brain tumor classification: Computational efficiency versus accuracy," 27th Annual Americas Conference on Information Systems, AMCIS 2021, ISBN: 978-173363258-4
- [20] Li. Daihui, Ma. Chengxu and Zeng. Shangyou, "Design of efficient convolutional neural module based on an improved module," Advances in Science, Technology and Engineering SystemsOpen AccessVolume 5, Issue 1, Pages 340 – 3452020,doi: 10.25046/aj050143
- [21] Pengyu. Li et al, "Improving CNN Model for Residential Building Image Classification: Enhancing Parameter Estimation Accuracy Through Transfer Learning and Reducing Model Complexity with MobileNet," Proceedings - 2023 3rd International Signal Processing, Communications and Engineering Management Conference, ISPCEM 2023Pages 50 – 542023, doi: 10.1109/ISPCEM60569.2023.00016
- [22] Vats. Satvik et al, "Advanced Image Classification on Intel Datasets Using Optimized EfficientNet and MobileNetV2," 2024 IEEE 9th International Conference for Convergence in Technology, I2CT 2024, doi: 10.1109/I2CT61223.2024.10543649
- [23] Zhou. Yingzi, Huang. Kun and Guo, Xiaoying, "End-to-end deep residual network for semantic segmentation," Journal of Physics: Conference SeriesOpen AccessVolume 1684, Issue 130 November 2020, doi: 10.1088/1742-6596/1684/1/012053

Improving Medical Image Segmentation Using Machine Learning: The Role of Optimization of Extracted Features in Convolutional Neural Networks

Moslem Kaviani, PhD Student ¹ | Omokolsoom Shahryari, Assistant Professor ²

¹Department of Computer Engineering, Sa.C, Islamic Azad University, Sanandaj, Iran.
mo.kaviani@iau.ac.ir

²Department of Computer Engineering, Sa.C, Islamic Azad University, Sanandaj, Iran.
shahryari.k@iausdj.ac.ir

Correspondence

Assistant Professor, Department of Computer Engineering, Sa.C, Islamic Azad University, Sanandaj, Iran.
Email: shahryari.k@iausdj.ac.ir

Received: - - 2025

Revised: 18 January 2024

Accepted: 12 February 2024

Abstract

In recent years, deep learning-based convolutional neural networks have made significant progress, especially in the field of medical image segmentation. However, factors such as uniform layer design, insufficient utilization of multi-scale information, and increased model complexity due to the large number of parameters have limited the performance of these methods in complex situations. In addition, data security and privacy are major challenges in medical data processing. In this study, an improved convolutional encoder-decoder model is introduced that uses strategies to optimize the extracted features and reduce the number of parameters. By employing basic information preservation mechanisms and dense attention modules, this model improves the ability to extract multi-level information and offers greater security and efficiency as an optimal option for use in a federated learning structure. Evaluation of the proposed model on the ClinicDB-CVC dataset shows that this method provides better performance in terms of intersection mean criteria on alliance and Dice coefficient compared to other state-of-the-art methods.

Keywords: Machine learning, medical image segmentation, federated learning, deep segmentation architecture with feature compression.