

مقایسه معماری خودرمزنگار پیچشی و ترنسفورمر بینایی در نویززدایی و بازسازی تصاویر سنگ های مخازن شکافدار زیرزمینی

بهراد تبریزی پور*

دانشکده مهندسی نفت، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

behrad.tabrizipour@srbiau.ac.ir

چکیده

سنگ های مخزنی شکافدار از مهم ترین اهداف در مطالعات مهندسی نفت به شمار می آیند، زیرا شبکه های پیچیده ی شکستگی ها و شکاف ها نقش تعیین کننده ای در تخلخل و نفوذپذیری این مخازن دارند. تصاویر سیتی اسکن به عنوان ابزاری کلیدی برای تحلیل ساختار درونی این سنگ ها استفاده می شوند، اما افت کیفیت تصاویر معمولاً بازسازی دقیق آن ها را دشوار می کند. در این پژوهش، دو معماری یادگیری عمیق شامل خودرمزگذار پیچشی و ترنسفورمر بینایی با هدف بهبود کیفیت و بازسازی تصاویر سیتی اسکن سنگ های مخزنی شکافدار طراحی و مقایسه شدند. وظیفه اصلی مدل ها بازسازی و نویززدایی تصاویر شکافدار زیرسطحی بود و پس از بازسازی، فرایند بخش بندی سه کلاسه شامل ماتریکس، شکاف باز و شکاف پر شده با استفاده از روش کلاسیک آستانه گذاری اتسو جهت بررسی عملکرد بهتر نتایج تصاویر بازسازی شده روی خروجی مدل ها اعمال شد. نتایج کمی و کیفی نشان داد که ترنسفورمر بینایی به واسطه بهره گیری از مکانیزم توجه، عملکرد برتری نسبت به خودرمزگذار دارد. این مدل در بازسازی تصاویر به نسبت سیگنال به نویز اوج ۴۵ و شاخص شباهت ساختاری ۰/۹۸ دست یافت، در حالی که خودرمزگذار به ترتیب مقادیر ۳۹ و ۰/۹۳ را ثبت کرد. در بخش بندی سه کلاسه مبتنی بر روش اتسو نیز ترنسفورمر بینایی در معیارهای دقت، بازخوانی و امتیاز F1 عملکرد بهتری داشت. برای کلاس ماتریکس، مقادیر دقت، بازخوانی و امتیاز F1 به ترتیب ۰/۹۸۹، ۰/۹۹۰ و ۰/۹۸۹ در مقایسه با ۰/۹۸۰، ۰/۹۶۹ و ۰/۹۷۴ برای خودرمزگذار به دست آمد. در کلاس شکاف پر شده، نتایج ۰/۹۵۳، ۰/۹۶۳ و در کلاس شکاف باز ۰/۹۲۳، ۰/۹۲۰ و ۰/۹۲۱ حاصل شد؛ در حالی که مدل خودرمزگذار مقادیر ۰/۹۶۰، ۰/۹۶۲، ۰/۹۶۱ و ۰/۹۴۰، ۰/۹۳۴، ۰/۹۳۷ را ثبت کرده بود. در مجموع، نتایج نشان می دهد که ترنسفورمر بینایی گزینه ای کارآمدتر و دقیق تر برای بازسازی تصاویر سیتی اسکن سنگ های شکافدار است و می تواند نقش مهمی در بهبود تحلیل های عددی خواص سنگ از جمله تخلخل و نفوذپذیری در مطالعات مهندسی مخزن ایفا کند.

کلمات کلیدی: سنگ مخزن شکافدار، بازسازی تصاویر، معماری خودرمزگذار، ترنسفورمر بینایی، نسبت سیگنال به نویز اوج، شاخص شباهت ساختاری

مقدمه

مخازن شکافدار به عنوان یکی از چالش‌های کلیدی مهندسی نفت مطرح هستند که مطالعه و درک آن‌ها اهمیت ویژه‌ای دارد. در تحقیقی، از شبکه عصبی مصنوعی (ANN) برای شناسایی نقاط شکافدار در مخازن کربناته استفاده شده که پیچیدگی توزیع ترک‌ها و نقش حیاتی آن‌ها در بهینه‌سازی استراتژی‌های حفاری و مدیریت مخزن را نشان می‌دهد [1]. همچنین، ولسون و همکاران به طور جامع به بررسی شناسایی و مدل‌سازی ترک‌ها در مخازن شکافدار طبیعی پرداخته‌اند و تأثیر این ترک‌ها بر افزایش بازیابی نفت را برجسته کرده‌اند [2]. سان و همکارش نیز بر اهمیت این مخازن به عنوان منابع اصلی تأمین انرژی جهانی تأکید دارند و به توسعه طبقه‌بندی‌های دقیق‌تر برای مدیریت مؤثرتر این مخازن پرداخته‌اند [3]. در مطالعه‌ای از لیو و همکاران، نقش ترک‌ها و شکستگی‌ها به عنوان کانال‌های نفوذ و فضای ذخیره‌سازی در سنگ‌های ماسه‌ای فشرده مورد بررسی قرار گرفته که این ویژگی‌ها بر غنی‌سازی و تولید نفت اثرگذار است [4]. در نهایت، هاشمی و همکاران و الربوی ویژگی‌های منحصر به فرد مخازن شکافدار، شامل تخلخل ماتریکس و هدایت ترک‌ها را تحلیل کرده و اهمیت آن‌ها را در جریان سیالات و بهینه‌سازی تولید بیان کرده‌اند. این تحقیقات نشان‌دهنده اهمیت شناخت و مدیریت دقیق مخازن شکافدار در توسعه بهینه منابع نفت و گاز هستند [5-6].

داده‌های تصویری در توصیف دقیق سنگ‌های مخزن اهمیت ویژه‌ای دارند، زیرا امکان بررسی ریزساختار، تخلخل و شکستگی‌ها را به صورت غیرمخرب و با جزئیات بالا فراهم می‌کنند. تکنولوژی سیتی اسکن یکی از ابزارهای اصلی و پیشرفته در این حوزه است که توانسته است با ارائه تصویربرداری سه‌بعدی با وضوح بالا، اطلاعات دقیقی از ساختار داخلی سنگ مخزن به دست آورد. ژو و همکاران نشان می‌دهند که سیتی اسکن نسبت به روش‌های سنتی که معمولاً مستلزم نمونه‌برداری مخرب و تحلیل‌های مستقیم هستند، مزیت‌های فراوانی دارد، از جمله امکان تصویربرداری غیرمخرب، بررسی دقیق‌تر توزیع تخلخل و تحلیل رفتار جریان سیالات در داخل منافذ و ترک‌ها [7]. از سوی دیگر، فیسال و همکاران نشان داده‌اند که سیتی اسکن می‌تواند به صورت کمی ساختارهای لایه‌های گل حفاری و آسیب‌های مخزن را نیز بررسی کند و با مزیت روش‌های غیرمخرب در مقایسه با تحلیل‌های کلاسیک، اطلاعات ارزشمندی ارائه دهد [8]. کاپوچینو و همکاران نیز با تأکید بر تصویربرداری سه‌بعدی سیتی اسکن، مزایای این ابزار در مطالعه تخلخل ماتریکس و شکستگی‌ها را برجسته ساخته‌اند که در روش‌های

مستقیم و مخرب قابل دسترسی نیست [9]. این مطالعات نشان می‌دهند که سیتی اسکن نقش بسیار مهمی در افزایش دقت و کیفیت مطالعات مخازن نفت ایفا می‌کند و تبدیل به ابزار ضروری در علم سنگ شده است.

این مطالعات به صورت جامع چالش‌های موجود در استفاده از داده‌های سیتی اسکن برای مطالعه مخازن نفت را بررسی می‌کنند و راهکارهایی برای بهبود کیفیت تصاویر ارائه می‌دهند که می‌تواند در بهبود دقت تحلیل خواص سنگ‌ها و مدل‌سازی مخازن حیاتی باشد.

داده‌های تصویربرداری سیتی اسکن با وجود مزایای فراوانی که ارائه می‌دهند، چالش‌های مهمی نیز همراه دارند که باید به آنها توجه شود. یکی از اصلی‌ترین مشکلات، نویز بالای تصاویر است که به دلیل پراکندگی اشعه ایکس و محدودیت‌های رزولوشن دستگاه‌های سیتی اسکن رخ می‌دهد، این مسئله منجر به کاهش کیفیت تصاویر و خطا در تفسیر دقیق خواص سنگ‌ها می‌شود. سلطان محمدی و همکارش این موضوع را به صورت جامع بررسی کرده و تأکید دارند که نویز ناشی از شرایط تصویربرداری می‌تواند دقت تحلیل‌های پتروفیزیکی را به شدت کاهش دهد [10]. برای حل این مشکل، روش‌های پیشرفته‌ای برای نویززدایی و بازسازی تصویر توسعه یافته‌اند که در تحقیقات خود به معرفی تکنیک‌های نوین بازسازی تصویر پرداخته و بهبود وضوح تصاویر و کاهش خطاهای تفسیر داده‌ها را نشان داده‌اند. همچنین بیژنی و همکاران اهمیت استفاده از الگوریتم‌های یادگیری عمیق جهت افزایش کیفیت داده‌های سیتی اسکن و کاهش نویز را برجسته کرده‌اند که این امر برای تحلیل‌های دقیق‌تر و مدل‌سازی بهتر مخازن حیاتی است [11]. این یافته‌ها بر اهمیت شناسایی دقیق ساختار شکستگی‌ها و تحلیل صحیح رفتار سنگ برای پیش‌بینی واکنش‌های سیال-سنگ در مخازن شکافدار تأکید دارند. این مطالعات نشان می‌دهند که با وجود چالش‌های موجود، استفاده از روش‌هایی که در پردازش تصاویر سیتی اسکن وجود دارند می‌تواند کیفیت داده‌های نهایی را بهبود بخشد و تفسیر درست خواص سنگ را تسهیل کند.

روش‌های کلاسیک نویززدایی مانند فیلترهای گاوسی، میانه و موجک اگرچه در تصویربرداری سیتی اسکن رایج هستند، اما مطالعات نشان داده‌اند که این روش‌ها در حفظ جزئیات ظریف با محدودیت جدی مواجه‌اند. شاهمرادی و همکاران در بررسی تجربی خود نشان دادند که فیلترهای کلاسیک با وجود کاهش نویز، موجب محوشدگی لبه‌ها و از دست رفتن بافت‌های ریز تصویر می‌شوند، در حالی که روش‌های پیشرفته‌تری مانند BM3D قادر به حذف نویز بدون آسیب‌زدن به ساختارهای بافتی مهم هستند [12]. فنگ و همکاران نیز در زمینه بازسازی تصاویر نویزی سیتی اسکن، گزارش کردند که فیلترهای متداول در

مواجهه با ساختارهای پیچیده و ناهمگن سنگ‌های مخزنی عملکرد ضعیفی دارند و در حفظ هندسه شکستگی‌ها ناکام می‌مانند [13]. علاوه بر این، مطالعات اخیر مانند پژوهش ساهو و همکاران که روش‌های ترکیبی موجک و شبکه‌های عمیق سه‌بعدی را بررسی کرده‌اند، نشان می‌دهد ترکیب فیلترهای موجک با معماری‌های پیچشی می‌تواند در تشخیص و بازسازی ترک‌های ریز در تصاویر میکروسیتی عملکرد بهتری ارائه دهد. با این حال، حتی این روش‌های کلاسیک-ترکیبی نیز در محیط‌های بسیار پیچیده محدودیت دارند [14]. به عنوان نمونه، متیو و همکاران در تصویربرداری میکروسیتی جریان‌های چندفازی مخازن متخلخل نشان دادند که روش‌هایی مانند NLM و انتشار ناهمسانگرد علی‌رغم بهبود سیگنال، نمی‌توانند پیوستگی منافذ ریز و ویژگی‌های چندمقیاسی را به درستی حفظ کنند [15]. این یافته‌ها تأکید می‌کنند که در مواجهه با ساختارهای پیچیده و نویزی سنگ‌های شکافدار، استفاده از روش‌های پیشرفته‌تر و مبتنی بر یادگیری عمیق ضروری است و روش‌های کلاسیک به تنهایی پاسخگوی نیازهای بازسازی دقیق تصاویر نیستند. روش‌های یادگیری عمیق تحول عظیمی در پردازش تصویر ایجاد کرده‌اند و به ویژه در تصویربرداری پزشکی و تحلیل ساختارهای پیچیده مانند مخازن نفت کاربردهای گسترده‌ای یافته‌اند. کوئزیر همکاران اصول فنی بازسازی تصویر سیتی اسکن با یادگیری عمیق را بررسی کرده و نشان دادند که این روش‌ها نسبت به تکنیک‌های سنتی، در حفظ جزئیات تصویر دقت بالاتری دارند [16]. همچنین جانسس و همکاران با بهره‌گیری از شبکه‌های مولد متخاصم موفق به افزایش وضوح تصاویر سیتی اسکن و بهبود بازسازی سه‌بعدی شده‌اند که در تحلیل ساختارهای پیچیده مخازن نقش بسزایی دارد [17].

در سال‌های اخیر، کاربردهای خاص در تحلیل سنگ‌ها و نمونه‌های متخلخل با استفاده از شبکه‌های عصبی گراف و تکنیک‌های مبتنی بر یادگیری انتقال‌پذیر توسعه یافته است. این تحقیقات به طور خاص در تشخیص شکستگی‌ها و ساختارهای ریز در سنگ‌های شکافدار و ناپایدار، با تمرکز بر دقت و سرعت بالا، موفق بوده‌اند. با این حال، محدودیت اصلی این مطالعات این است که اغلب بر سنگ‌های همگن و غیرشکافدار تمرکز دارند و کمتر به ساختارهای پیچیده و ناهمگن در مخازن شکافدار توجه شده است [18]. در حوزه تحلیل مخازن شکافدار، فناوری‌های پیشرفته‌تر مانند شبکه‌های مولد متخاصم و روش‌های ژنراتیو، نتایج قابل توجهی در افزایش وضوح و کاهش نویز ارائه داده‌اند که در بعضی موارد توانسته‌اند عملکرد برتری نسبت به روش‌های سنتی نشان دهند اما هنوز در مواجهه با ساختارهای پیچیده و ناپایدار، دقت کامل حاصل نشده است [19]. در نهایت، همزمان با

پیشرفت‌های فناوری، چالش‌های مربوط به دقت، پیچیدگی ساختارهای ناپایدار و محدودیت منابع داده از جمله تصاویر مخازن شکاف دار همچنان باقی است و نیازمند توسعه روش‌های ترکیبی و نوآورانه در تحلیل‌های تصویری مخازن شکافدار است.

با وجود پیشرفت‌های چشمگیر در پردازش تصاویر سیتی‌اسکن، بازسازی دقیق ساختارهای پیچیده در مخازن شکافدار همچنان با محدودیت‌هایی مواجه است. روش‌های کلاسیک نوین‌سازی نظیر فیلترهای گاوسی، میانه و موجک علی‌رغم کارایی اولیه در کاهش نویز، توان کافی در حفظ جزئیات حیاتی مانند ریزترک‌ها، مرزهای ماتریکس-شکاف و اتصال شکستگی‌ها را ندارند. از سوی دیگر، اغلب رویکردهای مبتنی بر یادگیری عمیق موجود، عمدتاً بر نمونه‌های همگن و بدون شکستگی متمرکز بوده و در مواجهه با داده‌های پیچیده، نویزی و ناهمگن حاصل از تصویربرداری سنگ‌های شکافدار، عملکرد مطلوبی در بازسازی ساختارهای ناپیوسته و حفظ ویژگی‌های فضایی ارائه نمی‌دهند. در این تحقیق، با بهره‌گیری از روش‌های مرتبط با یادگیری عمیق، چارچوبی نوین برای بازسازی تصاویر سیتی‌اسکن توسعه داده شده است که ضمن کاهش نویز، قادر به حفظ دقیق جهت‌گیری، پیوستگی و هندسه شکستگی‌ها می‌باشد. طراحی بهینه معماری مدل، انتخاب مناسب ابر پارامترها باعث شده است تا خروجی مدل به‌طور مستقیم قابل استفاده در تحلیل‌های پتروفیزیکی از جمله شبیه‌سازی نفوذپذیری، تخلخل دوگانه و جریان سیال در محیط‌های مخزنی پیچیده باشد. منظور از تخلخل دوگانه در این پژوهش، تخلخل مخازن شکافدار با دو مؤلفه‌ی مجزای تخلخل ماتریکس (ϕ_m) و تخلخل شکستگی (ϕ_f) است؛ به‌طوری‌که ماتریکس نقش اصلی در ذخیره‌سازی سیال و شکستگی‌ها نقش غالب در انتقال سیال را دارد و برهم‌کنش این دو محیط رفتار جریان را تعیین می‌کند. این رویکرد، گامی مؤثر در بهبود کیفیت داده‌های تصویری و ارتقاء دقت مدل‌سازی دیجیتال سنگ محسوب می‌شود.

جدول ۱ مروری ساختاریافته بر روش‌ها و معماری‌های رایج در پردازش تصاویر سنگ ارائه می‌دهد. همان‌طور که مشاهده می‌شود، روش‌های کلاسیک با وجود سرعت بالا، در حفظ ریزشکاف‌ها محدودیت دارند، در حالی‌که معماری‌های یادگیری عمیق به‌ویژه مدل‌های مبتنی بر توجه، توان بالاتری در بازسازی ساختارهای پیچیده و ناپیوسته شکستگی‌ها نشان داده‌اند. بر این اساس، در این پژوهش دو معماری CAE و ViT برای ارزیابی نقش معماری در کیفیت بازسازی شکستگی‌ها مقایسه شده‌اند.

جدول ۱ مروری بر مطالعات گذشته بر روی تصاویر سنگ جهت بازسازی یا نویزدایی تصویر

روش/معماری	نویسنده/نویسندگان شاخص	کاربرد اصلی در تصاویر سیتی اسکن سنگ	نقاط قوت	محدودیت‌ها	دامنه کاربرد/مقیاس مناسب
روش‌های کلاسیک محلی (موجک-گاوسی و میانه)	شاهمرادی و همکاران	نویزدایی اولیه	ساده، سریع، بدون نیاز به آموزش	محوشدن لبه‌ها و حذف ریزترک‌ها/بافت‌ها، ضعف در ناهمگنی	پیش پردازش سریع، نویز کم تا متوسط
روش‌های غیرمحلی کلاسیک NLM و مشابه	متیو و همکاران	نویزدایی با حفظ بافت	بافت نسبت به فیلترهای محلی	حساس به پارامترها، کند در D3، افت در نویز شدید	نویز متوسط، ساختارهای تکرارشونده منافذ/شکاف‌ها
BM3D / Non-local transform	شاهمرادی و همکاران	نویزدایی پیشرفته	حفظ جزئیات ریز و لبه‌ها بهتر از روش‌های محلی	محاسبات سنگین‌تر، محدودیت در نویزهای ساختاری پیچیده	تصاویر سیتی اسکن با نویز متوسط تا بالا، وقتی حفظ ریزشکاف‌ها مهم است
CNN های نویزدایی/بازسازی و DnCNN/ResNet/CAE و مشابه	Haeri, Bizhani, Little, Ardakani	نویزدایی و بازسازی	الگوی نویز، سرعت inference بالا، حفظ مناسب لبه شکاف‌ها	نیاز به داده آموزشی؛ احتمال oversmooth در ساختارهای شکستگی پراکنده	پرکاربرد در فیزیک سنگ دیجیتال و تصاویر مناسب ویژگی‌های محلی
شبکه‌های ابر تفکیک CNN/Residual (SR-CNN/VDSR/Residual Attention SR)	مرور و کاربرد برای تصاویر میکروسیتی Murugesu و همکاران	ابرتفکیک و بازیابی جزئیات زیررزولوشن	افزایش وضوح و دقت مرز منافذ/شکاف‌ها	نیاز به داده HR/LR یا طراحی unsupervised پایدار	مناسب حل trade-off FOV-resolution در تصاویر سیتی اسکن

سنگ‌های ناهمگن/شکافدار با شبکه ترک پیچیده	مدل‌سازی وابستگی‌های دوربرد، بهتر در ساختارهای چندمقیاسی ناپیوسته	بخش‌بندی دقیق ساختارهای چندمقیاسی	میکروسیتی زغال‌سنگ Chen نو: همکاران مرورهای جدید Brondolo نو: همکاران	ترنسفورمرهای بخش‌بندی Swin- T / TransUNet و ترکیبی‌ها
سنگ‌های پیچیده و نویزی، وقتی حفظ بافت حیاتی است	ریسک hallucination، آموزش ناپایدار، هزینه محاسباتی	بازیابی فرکانس‌های بالا، بهبود اتصال منافذ/شکاف‌ها	ابر تفکیک و بازسازی بافت/ریزساختار	GAN ها SRGAN ها برای فیزیک سنگ دیجیتال Huysmans, Janssens Swennen

در راستای رفع شکاف‌های موجود در ادبیات و با توجه به چالش‌های مطرح‌شده، این پژوهش مجموعه‌ای از نوآوری‌ها و مشارکت‌های علمی مشخص را ارائه می‌دهد که در ادامه به صورت خلاصه بیان می‌شوند:

(۱) یک چارچوب جدید مبتنی بر یادگیری عمیق برای نوپرزدایی و بازسازی تصاویر سنگ‌های مخازن شکافدار، با هدف حفظ هندسه و پیوستگی شکستگی‌ها.

(۲) مقایسه جامع دو معماری خودرمزگذار پیچشی و ترنسفورمر بینایی برای ارزیابی دقیق تأثیر معماری مدل بر کیفیت بازسازی.

(۳) به کارگیری معیارهای استاندارد نسبت سیگنال به نویز اوج، شاخص شباهت ساختاری، برای ارزیابی کمی، که نشان‌دهنده برتری ترنسفورمر بینایی در حفظ جزئیات ریز است.

(۴) استفاده از روش اتسو برای بخش‌بندی تصاویر بازسازی شده و واقعی جهت ارزیابی بیشتر تصاویر بازسازی شده نسبت به تصاویر واقعی به کمک معیارهایی نظیر شاخص جاکارد، امتیاز F1 و غیره.

روش کار

پیش‌پردازش داده‌ها

داده‌های خام این پژوهش از پایگاه Digital Rock Portal و توسط اسپینوزا و همکاران تهیه شده‌اند که شامل تصاویر سی‌تی اسکن از نمونه‌های مخازن شکاف‌دار زغال‌سنگی هستند (آدرس اینترنتی: <https://digitalporousmedia.org/published-datasets/drpfproject.published.DRP-21>) [20]. برای افزایش تعداد داده‌ها و تسهیل تحلیل محلی، تصاویر به قطعات 128×128 پیکسلی تقسیم شدند؛ پس از دریافت داده‌های خام، مجموع 2638 تصویر مورد استفاده قرار گرفت و مطابق روال اصلی، پیش از هرگونه پردازش و برای جلوگیری از نشت اطلاعات، با نسبت $10/10/80$ به مجموعه‌های آموزش، اعتبارسنجی و آزمون تقسیم شد. اسلایس‌های هر بخش سپس با یک اسکریپت اختصاصی به تصاویر 128×128 پیکسل تفکیک شدند. اندازه تصویر به صورت هدفمند انتخاب شد؛ به قدری بزرگ که اطلاعات بافتی و جزئیات لازم برای تشخیص شکستگی را حفظ کند، و در عین حال به اندازه‌ای کوچک که زمان محاسباتی در حد قابل قبول باقی بماند. این تفکیک بیشتر، حجم داده را به بیش از 11200 تصویر افزایش داد. در مرحله پیش پردازش، تصاویر فاقد شکستگی و شامل تنها ماتریکس حذف شدند؛ تعداد این تصویرها حدود 2400 مورد، معادل تقریباً 21 درصد از کل تصاویر اولیه بود. پس از حذف این نمونه‌ها، حدود 8800 تصویر باقی ماند. برای ایجاد توازن در نمایش دو نوع شکستگی «شکاف باز» و «شکاف پر شده»، تنها تصاویری انتخاب شدند که نسبت حضور این دو الگو در آنها یکنواخت بود؛ در نتیجه همان 8800 تصویر نهایی (حدود 78 درصد تصاویر اولیه) برای مراحل آموزش، اعتبارسنجی و آزمون مورد استفاده قرار گرفت. این داده‌ها در نهایت بر اساس نسبت 80 ، 10 و 10 درصد به ترتیب به 7040 نمونه آموزشی، 880 نمونه اعتبارسنجی و 880 نمونه آزمون تقسیم شدند. به منظور افزایش حجم داده‌ها و بهبود تعمیم‌پذیری مدل، از مجموعه‌ای از تبدیل‌های داده‌افزایی مبتنی بر تبدیلات هندسی استفاده شد. برای هر تصویر اولیه، سه نمونه‌ی افزوده‌شده تولید گردید که شامل وارونگی افقی تصادفی، وارونگی عمودی تصادفی و چرخش تصادفی در بازه‌ی $\pm 15^\circ$ درجه بود. علاوه بر این، برش مرکزی به عنوان مرحله‌ی پیش‌پردازش بر روی تمام تصاویر (اصلی و افزوده‌شده) اعمال شد تا ناحیه‌ی مرکزی تصویر با ابعاد ثابت (64) استخراج شود، بدون آن‌که خود این تبدیل منجر به افزایش تعداد نمونه‌ها گردد. در نتیجه، تعداد تصاویر آموزشی از 7040 تصویر اولیه (قبل از داده‌افزایی) به 28160 تصویر پس از داده‌افزایی افزایش یافت که معادل یک افزایش تقریباً 4 برابری در حجم داده است (جدول ۲). از آنجا که این تبدیل‌ها هندسه و توپولوژی شکستگی‌ها را حفظ می‌کنند، کیفیت ساختاری داده‌های افزوده‌شده با داده‌های اصلی هم‌راستا باقی می‌ماند. برای اطمینان از کیفیت، بخشی از داده‌های افزوده‌شده به صورت تصادفی بازبینی شد؛ نتایج نشان داد که مرز شکستگی‌ها، پیوستگی شبکه‌ی شکستگی‌ها و توزیع شدت پیکسل‌ها در تصاویر افزوده‌شده پایدار بوده و هیستوگرام آن‌ها تفاوت معناداری با داده‌های اصلی ندارد. این روش‌ها تنوع هندسی داده‌ها را افزایش

داده و محدودیت تعداد نمونه‌های واقعی (مانند نمونه‌های مغزه) را جبران کردند، که در فیزیک سنگ دیجیتال برای شبیه‌سازی خواص فیزیکی مانند نفوذپذیری حیاتی است. داده‌های افزوده‌شده به تنسورهای پایتورچ تبدیل و نرمال‌سازی شدند تا توزیع مناسب برای ورودی مدل‌های یادگیری عمیق فراهم شود. این نرمال‌سازی، پایداری گرادیان‌ها را در آموزش تضمین کرده و همگرایی را تسریع می‌کند، که برای پردازش تصاویر سیتی‌اسکن با کنتراست بالا در فیزیک سنگ دیجیتال ضروری است. همچنین لازم به ذکر است که نحوه نرمال‌سازی بر اساس میانگین و انحراف معیار واقعی داده‌های آموزش تعیین شده است؛ به این معنا که مقادیر آماری مجموعه داده ابتدا محاسبه شده و سپس روش نرمال‌سازی به گونه‌ای انتخاب شده که با توزیع شدت‌های تصاویر سازگار باشد. این رویکرد داده‌محور، توزیع ورودی یکنواخت‌تری ایجاد کرده و موجب پایداری عددی بیشتر در فرآیند آموزش مدل شد.

جدول ۲ تعداد داده‌های آموزش بعد از فرایند داده افزایی

نوع داده‌افزایی	تعداد کل افزوده‌شده	توضیح کیفیت / اثر
وارونگی افقی تصادفی	۷۰۴۰	حفظ هندسه شکستگی، افزایش تنوع الگوهای افقی، بدون تغییر توپولوژی
وارونگی عمودی تصادفی	۷۰۴۰	افزایش تنوع مکانی شکستگی‌ها، حفظ مرزها و پیوستگی ساختاری
چرخش تصادفی در بازه $\pm 15^\circ$ درجه	۷۰۴۰	افزایش تنوع جهت‌گیری، بدون تخریب ساختار ریزشکاف‌ها و مرزبندی‌ها
جمع کل داده‌های افزوده‌شده	۲۱۱۲۰	افزایش حجم داده
تعداد نهایی پس از داده‌افزایی	۲۸۱۶۰	معادل تقریباً ۴ برابر تعداد اولیه (۷۰۴۰ تصویر)

بازسازی تصاویر با استفاده از خودرمزگذار و ترنسفورمر بینایی

برای بازسازی تصاویر سیتی سنگ‌های مخازن شکاف‌دار، دو معماری یادگیری عمیق شامل خودرمزگذار پیچشی و ترنسفورمر بینایی طراحی و ارزیابی شدند. داده‌های ورودی شامل برش‌های سیتی اسکن با وضوح بالا بودند که ساختارهای پیچیده‌ای مانند ریزشکاف‌ها، شکستگی‌های متصل و نواحی ناهمگن ماتریکسی را نمایش می‌دادند. هدف اصلی مدل‌ها حذف نویز و تولید تصاویر

بازسازی شده‌ای بود که هندسه، جهت‌گیری و پیوستگی شکستگی‌ها را بدون تخریب حفظ کنند. نتایج کمی و کیفی نشان داد که ترنسفورمر بینایی در حفظ جزئیات ریز و ساختارهای چندمقیاسی عملکرد بهتری نسبت به خودرمزگذار دارد و می‌تواند داده‌های قابل اعتمادتری برای مراحل بعدی تحلیل دیجیتال سنگ، مانند شبیه‌سازی نفوذپذیری و جریان سیال، فراهم کند.

بخش‌بندی پایین‌دستی شکاف‌ها با آستانه‌گذاری اتسو

در این مطالعه، به منظور ارزیابی کیفیت بازسازی و میزان حفظ ساختارهای شکستگی، از روش آستانه‌گذاری چندسطحی اتسو به عنوان یک ابزار بخش‌بندی پایین‌دستی استفاده شد. این روش که مبتنی بر تحلیل هیستوگرام است و نیازی به آموزش ندارد، ماسک‌های سه کلاسه شامل ماتریکس، شکاف باز و شکاف پرشده را تولید می‌کند تا امکان بررسی تفکیک‌پذیری و حفظ مرزهای شکستگی پس از بازسازی فراهم شود. معیارهایی مانند جاکارد، امتیاز F1 بر اساس این ماسک‌ها محاسبه شدند و نقش آن‌ها تنها سنجش کیفیت بازسازی بوده است، زیرا بخش‌بندی هدف اصلی مدل‌ها نبوده و صرفاً برای ارزیابی نتایج بازسازی مورد استفاده قرار گرفته است.

در این پژوهش، هدف اصلی مدل‌ها انجام بازسازی و نویرزدایی تصاویر سنگ‌های مخازن شکاف‌دار است و بخش‌بندی نقش آموزشی یا مدل‌سازی ندارد. از آنجا که بخش‌بندی تنها به عنوان یک مرحله پایین‌دستی و برای ارزیابی میزان حفظ ساختارها پس از بازسازی استفاده شده است، به کارگیری یک معماری پیچیده و آموزش‌محور مانند U-Net که برای وظایف بخش‌بندی کاملاً نظارت‌شده طراحی شده است، در چارچوب مسئله حاضر ضرورت نداشته است. به همین دلیل، از روش کلاسیک آستانه‌گذاری چندسطحی اتسو استفاده شد که بدون نیاز به آموزش، امکان تولید ماسک‌های سه کلاسه (ماتریکس، شکاف باز و شکاف پرشده) را فراهم می‌کند و برای سنجش تفکیک‌پذیری ساختارها در تصاویر بازسازی شده کفایت می‌کند. در این پژوهش، تفکیک بین «شکاف باز» و «شکاف پرشده» در مرحله آموزش مدل‌ها انجام نمی‌شود و مدل‌ها صرفاً بازسازی تصویر را به صورت بدون نظارت یاد می‌گیرند. ارزیابی تفاوت این دو نوع شکستگی از طریق بخش‌بندی پایین‌دستی انجام شد؛ بدین صورت که

آستانه‌گذاری چندسطحی اتسو سه کلاس ماتریکس، شکاف باز و شکاف پر شده را بر اساس سطوح شدت خاکستری از یکدیگر جدا می‌کند (شکاف‌های باز معمولاً نواحی کم شدت/تیره‌تر و شکاف‌های پر شده نواحی با شدت بالاتر نسبت به شکاف باز هستند). سپس معیارهای Precision، Recall، F1-score و Jaccard برای هر کلاس به صورت مستقل محاسبه شده و عملکرد مدل‌ها در بازسازی و حفظ مرزهای «شکاف باز» و «شکاف پر شده» به طور جداگانه مقایسه گردید (جدول‌های ۴ و ۵ در بخش نتایج). ماتریس سردرگمی نیز میزان اشتباه بین این دو کلاس را نشان می‌دهد.

بنابراین مقادیر مربوط معیارهای ارزیابی این قسمت در این پژوهش تنها برای ارزیابی کیفیت بازسازی و میزان حفظ جزئیات شکستگی‌ها استفاده شده‌اند و نشان‌دهنده عملکرد مدل در یک وظیفه مستقل بخش‌بندی نیستند.

معماری خودرمزگذار تحقیق

معماری پایه خودرمزگذار توسط هینتون و همکارانش در سال ۱۹۸۶ و به شکل گسترده تر در سال ۲۰۰۶ معرفی شد که آن را به عنوان روشی برای کاهش ابعاد بدون استفاده از برچسب پیشنهاد دادند [21-22]. خودرمزگذار شامل دو بخش اصلی رمزگذار و رمزگشا می‌باشد. رمزگذار از چندین لایه پیچشی با فیلترهای ۳، ۴ و ۱۲۸ (به ترتیب برای لایه‌های اول تا سوم) برای کاهش بعد تشکیل شده است تا ویژگی‌های محلی مانند لبه‌های شکاف‌ها و بافت ماتریکس استخراج شوند. این لایه‌ها با تابع فعال‌ساز ReLU و max-pooling با کرنل 2×2 همراه بودند تا نویز تصاویر مانند نویز ناشی از محدودیت‌های تصویربرداری سی‌تی اسکن فیلتر شود. رمزگشا از لایه‌های پیچشی جابجا شده با ساختار معکوس برای افزایش بعد استفاده کرد تا تصاویر بازسازی‌شده با رزولوشن اصلی (128×128 پیکسل) تولید شوند. این معماری برای بازسازی جزئیات محلی شکاف‌ها بسیار مناسب است، زیرا در فیزیک سنگ دیجیتال، دقت در مقیاس‌های کوچک برای شبیه‌سازی جریان سیال در شکاف‌ها حیاتی است. برای تنظیم دقیق، از تکنیک عادی سازی دسته ای در لایه‌های میانی استفاده شد تا پایداری آموزش بهبود یابد.

معماری ترنسفورمر بینایی تحقیق

در این روش، برای بازسازی دقیق تصاویر سی‌تی اسکن سنگ‌های شکاف‌دار از معماری ترنسفورمر بینایی استفاده شد. این معماری نخستین بار در سال ۲۰۲۰ توسط دوسوویتسکی و همکارانش معرفی شد [23]. تفاوت اصلی آن با معماری‌های پیچشی

رایج، استفاده از مکانیزم توجه برای مدل‌سازی وابستگی‌های غیرمحلی بین بخش‌های تصویر است که باعث می‌شود ساختارهای پیچیده‌تری مانند شکاف‌های متصل و ریزترک‌ها به‌درستی شناسایی و بازسازی شوند.

ساختار پایه معماری ترنسفورمر بینایی

فرآیند تحلیل تصویر در ترنسفورمر بینایی به جای کانولوشن، مبتنی بر تقسیم تصویر به تصاویر مجزا و بررسی روابط بین آن‌ها است. این ساختار شامل مراحل کلیدی زیر می‌باشد:

تقسیم تصویر به تصاویر ثابت

ابتدا تصویر سی‌تی اسکن با اندازه‌ی ثابت (مثلاً 128×128 پیکسل) به بخش‌های کوچک‌تر با اندازه مساوی (مثلاً 16×16 پیکسل) تقسیم می‌شود. هر تصویر دوبعدی به یک بردار ویژگی یک‌بعدی با طول ثابت تبدیل می‌شود. برای این کار، از یک لایه‌ی مسطح‌ساز برای تبدیل هر تصویر به یک بردار و سپس یک لایه‌ی کاملاً متصل برای نگاشت آن به فضای ویژگی استفاده می‌شود. معمولاً مقدار برابر با ۵۱۲ یا ۷۶۸ انتخاب می‌شود. این بردارها نشان‌دهنده‌ی ویژگی‌های آماری، بافتی و ساختاری هر تصویر هستند و پایه‌ای برای مدل‌سازی توجه بین تصویرها در مراحل بعدی فراهم می‌کنند. از آنجا که ترنسفورمر بینایی برخلاف شبکه عصبی پیچشی فاقد مکانیزم ضمنی برای درک موقعیت مکانی اجزای تصویر است، به هر بردار ویژگی یک بردار مکانی افزوده می‌شود تا موقعیت نسبی هر تصویر در تصویر حفظ گردد.

پردازش با لایه‌های رمزنگار

پس از آماده‌سازی توالی تصاویر کدگذاری‌شده، این بردارها به لایه‌های متعدد خودتوجهی چندوجهی و شبکه عصبی مصنوعی وارد می‌شوند. مکانیزم توجه سراسری در این لایه‌ها روابط بین تصاویر مختلف را بدون در نظر گرفتن فاصله مکانی، مدل‌سازی می‌کند. در این پژوهش، ۸ لایه رمزگذار با ۴ مکانیزم توجه که دارای ضرب داخلی، مقیاس پذیری، گذر از تابع فعال‌ساز softmax و سپس ضرب داخلی مجدد می‌باشد (خروجی این ۴ مکانیزم توجه با هم پیوند داده می‌شود) و بردار ویژگی با بعد ۵۱۲ استفاده شد.

بازسازی یا طبقه‌بندی نهایی با رمزگشا

خروجی نهایی معماری ترنسفورمر بینایی می‌تواند برای اهداف متنوعی از جمله بازسازی تصویر، طبقه‌بندی اجزا و تشخیص الگوها مورد استفاده قرار گیرد. لایه‌ی رمزگشا سه ورودی دریافت می‌کند؛ دو ورودی از سمت لایه‌ی رمزگذار و یک ورودی از خود رمزگشا. در این لایه نیز از مکانیزم توجه بهره گرفته می‌شود، با این تفاوت که پس از مرحله‌ی مقیاس‌گذاری، یک ماسک به آن افزوده می‌شود که این فرآیند با عنوان توجه ماسک‌دار شناخته می‌شود. در این مکانیزم، بخش بالایی مثلثی ماتریس توجه پس از عبور از تابع Softmax صفر می‌شود تا در عملیات ضرب برداری بعدی، مدل تنها به ورودی‌های جاری و پیشین تمرکز داشته باشد و از دسترسی به اطلاعات آینده، که از سمت رمزگذار تأمین می‌شوند، جلوگیری شود. زمانی که مدل به مرحله‌ی پردازش آن ورودی‌ها برسد، امکان توجه به تمامی ورودی‌ها از جمله ورودی‌های آینده نیز فراهم می‌گردد. در نهایت، با استفاده از لایه‌های رمزگشا، تصویر بازسازی‌شده‌ی نهایی تولید می‌شود.

مزایای ترنسفورمر بینایی در تحلیل تصاویر سیتی اسکن مخازن شکافدار

یکی از چالش‌های اصلی در تحلیل تصاویر سیتی اسکن سنگ‌های مخزن، پراکندگی شدید شکاف‌ها و نبود ساختار پیوسته در تصویر است. معماری ترنسفورمر بینایی با درک وابستگی‌های دوربرد قادر است الگوهای شکستگی را حتی اگر در فواصل زیاد از یکدیگر قرار داشته باشند، به خوبی شناسایی و بازسازی کند. همچنین، استفاده از بردارهای ویژگی برای هر تصویر، موجب می‌شود مدل نه تنها شکل هندسی، بلکه بافت و شدت خاکستری مناطق مختلف تصویر را نیز در نظر بگیرد.

معماری ترنسفورمر بینایی تحقیق

برای بازسازی دقیق تصاویر سیتی اسکن از سنگ‌های شکافدار، مدلی مبتنی بر معماری ترنسفورمر بینایی طراحی شد. هدف این مدل، حفظ و بازسازی ساختارهای پیچیده مانند شکاف‌های متصل و ریزترک‌ها در تصاویر نویزی و کم‌وضوح است.

در مرحله اول، تصویر ورودی با اندازه ۱۲۸ در ۱۲۸ پیکسل به تصاویری با ابعاد ۱۶ در ۱۶ تقسیم شد. هر تصویر به عنوان یک واحد مستقل، پس از مسطح‌سازی، به برداری با طول ۵۱۲ نگاشت شد تا اطلاعات محلی آن در قالب بردار ویژگی حفظ گردد.

این بردارها از طریق یک لایه خطی به فضای ویژگی منتقل شدند. سپس برای حفظ موقعیت مکانی هر تصویر در ساختار تصویر، بردارهای موقعیتی به توالی افزوده شد و مجموعه‌ی حاصل به عنوان ورودی به رمزگذار ترنسفورمر بینایی ارسال گردید.

رمزگذار شامل یک یا چند لایه توجه با چند سر توجه بوده که با هدف درک روابط بین تصویرها در تصویر، وابستگی‌های غیرمحلی را نیز مدل می‌کند. پس از رمزگذاری، بردارهای خروجی با میانگین‌گیری، فشرده‌سازی شده و به بردار نهفته‌ای در فضای ویژگی کاهش داده شدند. این بردار برای مرحله بازسازی به تعداد تصویرها تکرار شده و به‌عنوان ورودی دیکدر مورد استفاده قرار گرفت.

رمزگشا که ساختاری مشابه رمزگذار داشت، با استفاده از حافظه‌ی خروجی رمزگذار، توالی‌ای از بردارهای بازسازی‌شده تولید نمود. این بردارها سپس مجدداً به تصاویر تصویری بازگردانده شده و در موقعیت مکانی اولیه خود جای‌گذاری شدند تا تصویر نهایی با وضوح بالا بازسازی گردد. همچنین ابرپارامترهای بهینه مورد استفاده در این شبکه نیز در جدول ۱ و ۲ مشخص شده است. تصاویر بازسازی‌شده با رزولوشن 128×128 پیکسل به عنوان تصاویر دیجیتال مخزن برای شبیه‌سازی عددی و پیش‌بینی خواص فیزیکی مانند نفوذپذیری، تخلخل دوگانه، و جریان سیال استفاده شدند. این خروجی‌ها در فیزیک سنگ دیجیتال برای پیش‌بینی رفتار مخزن در سناریوهای واقعی مانند تزریق آب یا شکست هیدرولیکی کاربرد دارند.

مدل‌سازی

نحوه آموزش و آزمایش

برای آموزش معماری‌های خودرمزگذار و ترنسفورمر بینایی، داده‌های آماده و پیش پردازش شده به سه مجموعه تقسیم شدند: ۸۰٪ برای آموزش، ۱۰٪ برای اعتبارسنجی، و ۱۰٪ برای آزمایش. این تقسیم‌بندی به منظور اطمینان از تعادل بین یادگیری مدل، تنظیم ابر پارامترها، و ارزیابی عملکرد در داده‌های نادیده‌شده انجام شد. داده‌های آموزشی شامل تصاویر 128×128 پیکسل از مخازن شکافدار مانند سیتی اسکن بودند که با تکنیک‌های داده‌افزایی (مانند چرخش تصادفی و برش مرکزی) تقویت شده بودند تا تنوع کافی برای یادگیری الگوهای پیچیده شکاف‌ها فراهم شود. آموزش مدل‌ها با استفاده از کتابخانه پایتورچ و بهینه‌ساز آدام (با نرخ یادگیری اولیه 0.001) انجام شد. برای بهبود همگرایی و جلوگیری از بیش‌بردازش، از یک scheduler پویا برای تنظیم نرخ یادگیری که در این تحقیق از کاهش خطی برای هر دو مدل پس از ۵۰ اپوک استفاده شد.

نحوه بهبود مدل و بهینه‌سازی

در این مرحله، برای بهینه‌سازی عملکرد مدل روی داده‌های خاص مخازن شکاف‌دار، تنظیم دقیق روی مدل‌های آموزش‌دیده انجام شد. وزن‌های لایه‌های پیچیده در خودرمزگذار و ترنسفورمر بینایی با استفاده از ۱۰٪ داده‌های اعتبارسنجی به‌روزرسانی شدند تا مدل حساسیت بیشتری به ویژگی‌هایی مانند میکروشکاف‌ها و اتصال شبکه شکاف‌ها پیدا کند. برای جلوگیری از بیش‌برازش، از تکنیک دراپ‌اوت با نرخ ۰/۲ در لایه‌های میانی استفاده شد. این کار باعث افزایش تعمیم‌پذیری مدل در مواجهه با داده‌های محدود و متنوع پایگاه Digital Rock Portal شد.

تنظیم ابرپارامترها با استفاده از روش جستجوی شبکه انجام شد تا مقادیر بهینه برای پارامترهای کلیدی شامل نرخ یادگیری (بین ۰/۰۰۰۱ تا ۰/۱)، اندازه دسته (۱۶، ۳۲، ۶۴)، توابع هزینه (MSE و MAE)، و توابع بهینه‌ساز (Adam, AdamW, SGD, RMSProp) تعیین شوند. از آنجا که جستجوی شبکه به طور سنتی برای الگوریتم‌های یادگیری ماشین طراحی شده و در یادگیری عمیق به دلیل پیچیدگی محاسباتی مستقیماً قابل استفاده نیست، از یک ارتباط دهنده برای ارتباط دادن جستجوی شبکه با مدل‌های یادگیری عمیق استفاده شد. این ارتباط دهنده امکان جستجوی خودکار ابرپارامترها را در فضای بزرگ پارامترها فراهم کرد. فرآیند تنظیم با ارزیابی معیارهای عملکرد (مانند F1-score برای کلاس شکاف و شاخص شباهت ساختاری برای کیفیت بازسازی) روی مجموعه اعتبارسنجی انجام شد تا بهترین ترکیب ابرپارامترها انتخاب شود. علاوه بر این، از تکنیک توقف زودهنگام استفاده شد تا آموزش در صورت عدم بهبود عملکرد روی مجموعه اعتبارسنجی پس از ۱۰ اپوک متوقف شود، که زمان محاسبات را کاهش داد و از بیش‌برازش جلوگیری کرد. جدول ۳ و ۴ نشان‌دهنده پارامترهای بهینه استفاده شده در مدل سازی نهایی خودرمزگذار و ترنسفورمر بینایی میباشد.

جدول ۳ ابرپارامترهای بهینه بعد از استفاده جستجو شبکه در معماری خودرمزگذار

مقدار	ابر پارامتر
۸	بج سائز
MAE	تابع هزینه
Adam	بهینه ساز
(0.9, 0.999)	پارامترهای betas بهینه ساز
.	پارامتر Weight decay بهینه سازی
۰.۰۰۰۱	نرخ یادگیری

جدول ۴ ابرپارامترهای بهینه بعد از استفاده جستجو شبکه در معماری ترنسفورمر بینایی

مقدار	ابر پارامتر
۱۲۸	بج سائز
MSE	تابع هزینه
Adam	بهینه ساز
(0.9, 0.999)	پارامترهای betas بهینه ساز
۰	پارامتر Weight decay بهینه ساز
۰.۰۱	نرخ یادگیری
۸	تعداد لایه های رمزگذار
۴	تعداد اتشن در رمزگذار و رمزگشا
۵۱۲	بردار ویژگی هر تصویر

برای بهبود کارایی محاسباتی، آموزش روی GPU مدل NVIDIA RTX 3050 انجام شد تا پردازش تصاویر 128×128 پیکسل با حجم بالا تسریع شود. این سخت افزار امکان اجرای سریع تر آموزش و تنظیم مدل ها را فراهم کرد، که برای مدیریت داده های پیچیده مخازن شکافدار ضروری بود.

نحوه ارزیابی مدل روی داده های دیده نشده و آزمایش

ارزیابی با استفاده از معیارهای کمی شامل شاخص جاکارد، نسبت سیگنال به نویز اوج، شاخص شباهت ساختاری، و دیگر معیارهای ارزیابی پر کاربرد مدل های طبقه بندی انجام شد. دقت بخش بندی شکاف ها را در برابر ماتریکس سنجید، که به دلیل عدم تعادل ذاتی داده های مخازن شکافدار (کمتر از ۱۰٪ شکاف در تصاویر) وجود دارد. شاخص جاکارد، میزان هم پوشانی بین مناطق پیش بینی شده شکاف و مناطق واقعی را ارزیابی کرد. نسبت سیگنال به نویز اوج کیفیت بازسازی تصاویر را از نظر شباهت پیکسل به پیکسل با تصاویر مرجع سنجید، که برای اطمینان از حفظ جزئیات ساختاری (مانند لبه های شکاف) در فیزیک سنگ دیجیتال ضروری است. شاخص شباهت ساختاری با تمرکز بر حفظ ساختارهای فضایی و الگوهای شکاف (مانند اتصال شبکه های شکاف)، کیفیت بازسازی را از منظر شباهت ساختاری بررسی کرد. دیگر معیارهای ارزیابی مانند دقت، بازیابی، امتیاز F1، ماتریس

سردرگمی و شاخص جاکارد نیز در این تحقیق برای بررسی دقت بازیابی کلاس های متنوع ماتریس، شکاف پر شده و شکاف خالی بین تصاویر اصلی و تصاویر بازسازی شده مورد استفاده قرار گرفت.

ارزیابی کمی کیفیت بازسازی تصاویر با شاخص های شاخص نسبت سیگنال به نویز اوج، شاخص

شباهت ساختاری و معیارهای طبقه بندی

شاخص نسبت سیگنال به نویز اوج

نسبت سیگنال به نویز اوج یکی از متداول ترین معیارهای ارزیابی کیفیت بازسازی تصویر است که برای ارزیابی کیفیت تصویری در سیستم های فشرده سازی دیجیتال معرفی و به صورت گسترده در پردازش تصویر و یادگیری عمیق مورد استفاده قرار گرفت [24]. این شاخص میزان انحراف تصویر بازسازی شده از تصویر مرجع را برحسب دسی بل بیان می کند و به صورت زیر تعریف می شود:

$$PSNR = 10 * \log_{10}((MAX^2) / MSE) \quad (1)$$

$$MSE = (1 / (m * n)) * \sum_{i=1}^m \sum_{j=1}^n (I_r(i,j) - I_t(i,j))^2 \quad (2)$$

که در آن:

MAX: بیشترین مقدار ممکن برای یک پیکسل در تصویر. برای مثال در تصاویر ۸ بیتی خاکستری مقدار ۲۵۵ است. (چون پیکسل ها از ۰ تا ۲۵۵ مقدار میگیرند)

MSE: میانگین مربعات خطا (Mean Squared Error) بین تصویر مرجع و تصویر بازسازی شده. هر چه MSE کمتر باشد، خطا کمتر و کیفیت بالاتر است.

$I_r(i,j)$: مقدار پیکسل در سطر i و ستون j از تصویر واقعی

$I_i(i,j)$: مقدار پیکسل در همان موقعیت (i,j) از تصویر بازسازی شده یا فشرده شده

m : تعداد ستون های تصویر (عرض تصویر به تعداد پیکسل)

n : تعداد سطرهای تصویر (ارتفاع تصویر به تعداد پیکسل)

در واقع، هرچه مقدار نسبت سیگنال به نویز اوج بالاتر باشد، کیفیت بازسازی بهتر و میزان نویز در تصویر کمتر است.

شاخص شباهت ساختاری

شاخص شباهت ساختاری معرفی شد و به عنوان جایگزینی دقیق تر برای نسبت سیگنال به نویز اوج مطرح گردید، زیرا به جای تمرکز صرف بر اختلاف پیکسل ها، ساختار، روشنایی و کنتراست تصویر را نیز در نظر می گیرد [25]. شاخص شباهت ساختاری یکی از دقیق ترین معیارهای ارزیابی کیفیت تصویری به ویژه در کاربردهای پزشکی و مهندسی سنگ است.

فرمول اصلی شاخص شباهت ساختاری به صورت زیر تعریف می شود:

$$SSIM(x, y) = ((2 * \mu_x * \mu_y + C1) * (2 * \sigma_{xy} + C2)) / ((\mu_x^2 + \mu_y^2 + C1) * (\sigma_x^2 + \sigma_y^2 + C2)) \quad (3)$$

که در آن:

x و y : دو تصویر مورد مقایسه (یکی تصویر بازسازی شده، دیگری تصویر مرجع)

μ_x , μ_y : میانگین شدت پیکسل های دو تصویر x و y

σ_x^2 , σ_y^2 : واریانس شدت پیکسل های دو تصویر

σ_{xy} : کوواریانس بین دو تصویر (میزان همبستگی بین x و y)

C_1 , C_2 : ضرایب پایدارسازی برای جلوگیری از تقسیم بر صفر

L: بیشترین مقدار شدت پیکسل در تصویر (برای تصاویر ۸ بیتی معمولاً ۲۵۵)

مقدار شاخص شباهت ساختاری بین ۰ و ۱ قرار دارد؛ مقدار ۱ نشان‌دهنده‌ی شباهت کامل بین دو تصویر است. برخلاف نسبت سیگنال به نویز اوج که صرفاً بر اساس اختلاف پیکسل‌ها محاسبه می‌شود، شاخص شباهت ساختاری ساختارهای محلی تصویر را مقایسه کرده و حساسیت بیشتری نسبت به تغییرات بافت و روشنایی دارد.

دقت

شاخص دقت، نسبت پیکسل‌هایی را نشان می‌دهد که مدل به‌درستی به‌عنوان «مثبت» (مانند ناحیه‌ی شکاف) پیش‌بینی کرده است، نسبت به کل پیکسل‌هایی که مدل به‌عنوان مثبت تشخیص داده است [26]. مقدار بالای دقت بیانگر توانایی مدل در کاهش خطاهای مثبت کاذب (False Positive) است؛ یعنی مدل از برجسب‌گذاری اشتباه نواحی غیرشکاف به‌عنوان شکاف جلوگیری می‌کند. این ویژگی در کاربردهایی مانند شناسایی شکستگی‌ها اهمیت بالایی دارد، زیرا از «بیش‌برآورد نواحی شکاف‌دار» جلوگیری می‌کند. فرمول محاسبه‌ی دقت به صورت زیر است:

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

که در آن:

- TP (True Positive) تعداد پیکسل‌هایی است که مدل به‌درستی به‌عنوان شکاف (کلاس مثبت) پیش‌بینی کرده است.
- FP (False Positive) مواردی است که مدل به‌اشتباه پیکسل‌هایی را به‌عنوان شکاف شناسایی کرده در حالی که در واقع ماتریکس هستند.

بازخوانی یا یادآوری

شاخص بازخوانی نشان‌دهنده‌ی نسبت پیکسل‌های واقعی شکاف است که مدل آن‌ها را به‌درستی شناسایی کرده است. مقدار بالای این معیار به این معناست که مدل بیشتر شکاف‌های واقعی را تشخیص داده و نواحی شکستگی را از دست نداده است [27].

فرمول آن به شکل زیر است :

$$Recall = \frac{TP}{TP + FN}$$

که در آن:

FN (False Negative) پیکسل‌هایی هستند که به فرض مثال مدل به اشتباه آن‌ها را به عنوان ماتریکس شناسایی کرده در حالی که در واقع شکاف بوده‌اند.

در حوزه‌ی تحلیل تصاویر سی‌تی اسکن سنگ‌های مخازن، مقدار بازخوانی بالا از اهمیت ویژه‌ای برخوردار است، زیرا از نادیده گرفته شدن بخش‌هایی از شبکه‌ی شکستگی جلوگیری می‌کند.

امتیاز - F1

امتیاز F1 میانگین هارمونیک دو شاخص دقت و بازخوانی است و تعادلی میان خطاهای مثبت کاذب و منفی کاذب برقرار می‌کند. این معیار در داده‌هایی با عدم تعادل بین کلاس‌ها (مانند نسبت کم شکاف‌ها نسبت به ماتریکس) اهمیت زیادی دارد، زیرا از تأثیر نامتوازن هر یک از معیارها جلوگیری می‌کند. فرمول آن به صورت زیر بیان می‌شود:

$$F1\text{-score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

مقدار بالاتر F1 نشان‌دهنده‌ی عملکرد کلی بهتر مدل در تشخیص صحیح نواحی شکستگی و جلوگیری از خطاهای طبقه‌بندی است.

شاخص جاکارد

برای ارزیابی جامع‌تر عملکرد مدل، از شاخص جاکارد استفاده شد. این شاخص میزان هم‌پوشانی بین نواحی پیش‌بینی شده و نواحی واقعی را اندازه‌گیری کرده و به صورت زیر تعریف می‌شود [28]:

$$Jaccard\ Score = \frac{|A \cap B|}{|A \cup B|} \quad (7)$$

که در آن مجموعه‌ی A مجموعه‌ی مقادیر پیش‌بینی شده است و مجموعه‌ی B مجموعه‌ی مقادیر واقعی. عبارت $A \cup B$ شامل هر چیزی است که در هر یک از دو مجموعه‌ی A و B وجود دارد، و $A \cap B$ نشان‌دهنده‌ی بخش هم‌پوشان (اشتراک) بین دو کلاس است.

تمام معیارها برای هر کلاس به صورت جداگانه محاسبه شده‌اند و سپس میانگین گرفته شده تا هم امتیازهای ویژه‌ی هر کلاس و هم عملکرد کلی به دست آید. محاسبات با استفاده از کتابخانه‌های استاندارد و معتبر انجام شدند و شامل مقایسه‌ی پیکسل به پیکسل بین ماسک‌های سگمنتیشن پیش‌بینی شده و برچسب‌های واقعی در مجموعه داده‌ی آزمایشی بودند.

شاخص جاکارد بین ۰ تا ۱ تغییر می‌کند؛ هرچه مقدار آن به ۱ نزدیک‌تر باشد، تطابق بیشتری بین تصویر بازسازی شده و تصویر مرجع وجود دارد.

مقایسه جامع هزینه‌های محاسباتی و نیازهای زیرساختی دو معماری بازسازی تصویر

اگرچه معماری ترنسفورمر بینایی در اغلب معیارهای کمی و کیفی عملکرد دقیق‌تری نسبت به خودرمزگذار ارائه داد، اما بررسی جامع هزینه‌های محاسباتی نشان می‌دهد که انتخاب این معماری مستلزم برخورداری از زیرساخت‌های سخت‌افزاری قوی‌تر است. برای پردازش تصاویر 128×128 ، مدل خودرمزگذار به دلیل کانولوشن‌های محلی و تعداد پارامترهای کمتر، مصرف حافظه GPU در حدود ۱ تا ۱.۲ گیگابایت داشته و می‌تواند با بچ سایزهای بزرگ‌تری (تا ۶۴) بدون افت عملکرد اجرا شود. در مقابل، مدل ترنسفورمر بینایی به علت وجود لایه‌های توجه چندسری و تعبیه‌های با ابعاد ۵۱۲، بین ۲ تا ۲.۸ گیگابایت حافظه VRAM مصرف می‌کند. از نظر زمان آموزش، خودرمزگذار در هر اپوک بین ۱۸ تا ۲۰ ثانیه زمان نیاز دارد، در حالی که ترنسفورمر بینایی به طور متوسط ۳۲ تا ۳۸ ثانیه برای هر اپوک صرف می‌کند و برای رسیدن به همگرایی نیز به تعداد تکرارهای بیشتری نیازمند است، که مجموع زمان آموزش را ۱.۵ تا ۲ برابر افزایش می‌دهد. علاوه بر GPU، بررسی مصرف RAM نشان داد که خودرمزگذار با ۲ تا ۳ گیگابایت RAM به خوبی اجرا می‌شود، در حالی که ترنسفورمر بینایی به علت ذخیره تصاویر، تعبیه‌ها و توالی‌های توجه، نیاز به ۴ تا ۷ گیگابایت RAM دارد. همچنین تحلیل توان پردازشی نشان داد که در سراسر مدل خودرمزگذار حدود ۴۵۰ تا ۶۰۰ تصویر در ثانیه است، در حالی که در ترنسفورمر بینایی این مقدار به حدود ۱۸۰ تا ۲۳۰ تصویر در ثانیه کاهش می‌یابد. در نهایت، اگرچه ترنسفورمر بینایی در بازسازی ساختارهای پیچیده مخازن شکاف‌دار عملکرد برتری دارد، اما هزینه محاسباتی بالاتر،

سرعت پایین‌تر، نیاز به GPU با حداقل ۸ گیگابایت حافظه و وابستگی بیشتر به تنسور هسته‌ها باعث می‌شود که در پروژه‌های وسیع، تکرار شونده یا زیرساخت‌های محدود، مدل خودرمزگذار گزینه‌ای عملی‌تر و مقرون‌به‌صرفه‌تر باشد. انتخاب بین این دو معماری باید بر پایه سطح پیچیدگی داده‌ها، حجم کل داده‌ها، محدودیت GPU و الزام به دقت بازسازی انجام گیرد. در جدول ۵ نیز اطلاعات لازمه مشخص شده است.

جدول ۵ اطلاعات لازمه محاسباتی هردو مدل بازسازی تصویر

معیار	خودرمزگذار	ترنسفورمر بینایی
مدل GPU	RTX 3050	RTX 3050
حداقل GPU لازم	۴ GB کافی	حداقل ۸ GB برای اجرای پایدار
مصرف RAM سیستم	2-3 GB	4-7 GB
سرعت پردازش	450-600 تصویر/ثانیه	180-230 تصویر/ثانیه
زمان هر اپوک	18-20 ثانیه	32-38 ثانیه
تعداد اپوک تا همگرایی	25-30	40-50
حساسیت به افزایش بچ سائز	مقیاس‌پذیری عالی	محدودیت شدید OOM
پایداری آموزش	معمولی	بسیار بالا
توانایی یادگیری ویژگی‌های غیرمحلی	محدود	بسیار قوی
مناسب برای داده حجیم	تا حدی	عالی اما پرهزینه
کاربرد پیشنهادی	پروژه‌های کم‌منبع، آموزش سریع	پروژه‌های با دقت بالا و شکستگی‌های پیچیده
هزینه زمانی-پردازشی کل	کم	زیاد (1.5-2× AE)

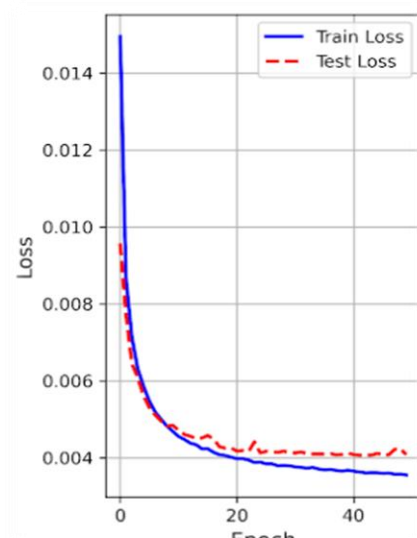
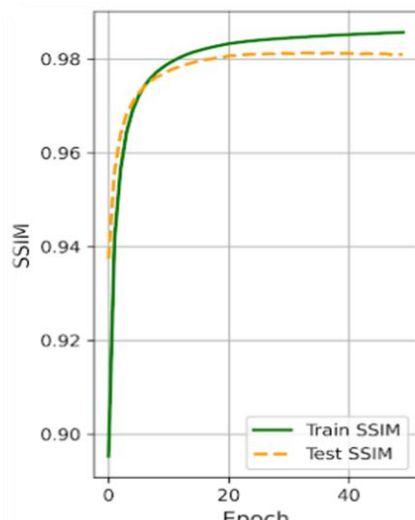
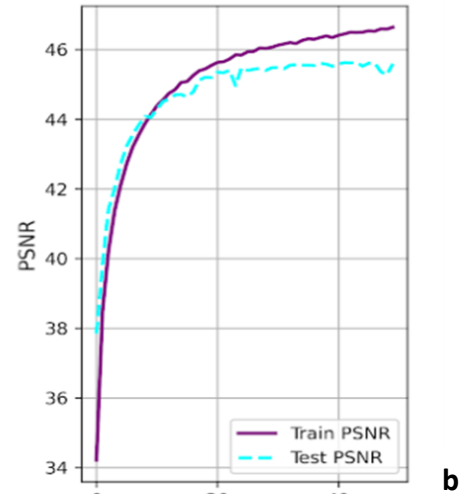
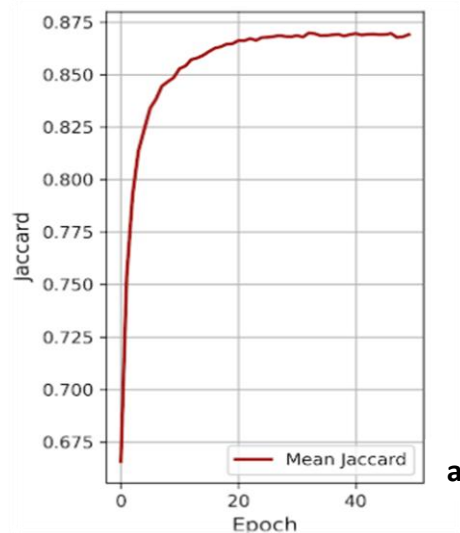
نتایج و بحث

در این پژوهش، دو معماری متفاوت از معماری‌های یادگیری عمیق برای بازسازی و نویززدایی تصاویر سی‌تی اسکن سنگ‌های مخازن شکافدار مورد مقایسه قرار گرفتند:

۱. معماری خودرمزگذار

۲. معماری ترنسفورمر بینایی

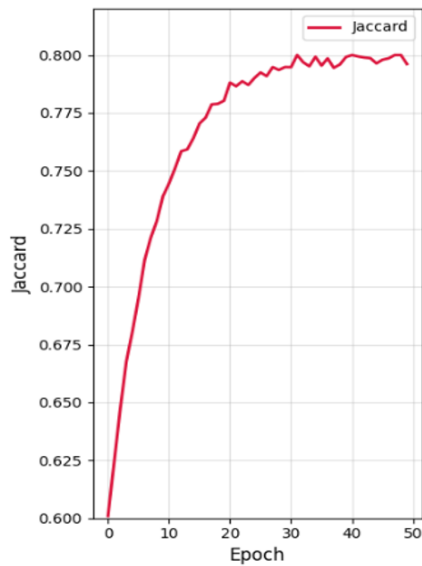
نتایج به دست آمده نشان می‌دهد که معماری ترنسفورمر بینایی در اغلب معیارها عملکرد بهتری نسبت به معماری خودرمزگذار داشته است. به ویژه در شاخص نسبت سیگنال به نویز اوج با عدد بیشتر از ۴۰ (شکل ۱b)، معماری ترنسفورمر بینایی توانست وضوح بالاتری در بازسازی تصویر ایجاد کند که نشان‌دهنده بازتولید دقیق‌تر شدت پیکسل‌ها و جزئیات ریز شکاف‌ها نسبت به مدل خودرمزگذار (شکل ۲b) است. همچنین، شاخص شباهت ساختاری برای ترنسفورمر بینایی مقادیر بالاتری (نزدیک تر به ۱) (شکل ۱c) نسبت به شاخص شباهت ساختاری برای مدل خودرمزگذار (شکل ۲c) را ثبت کرده است، که بیانگر حفظ بهتر ساختار فضایی و همبستگی‌های موضعی در تصویر می‌باشد. شاخص جاکارد نیز نشان داده است که مدل ترنسفورمر توانایی بیشتری در تفکیک دقیق نواحی شکاف باز و ماتریکس سنگ داشته است (شکل ۱a و ۲a). در مقابل، معماری خودرمزگذار در بازسازی بخش‌های همگن عملکرد قابل قبولی داشته اما در بازتولید تصاویر و شکاف‌ها مقداری نویز در بر گرفته است.



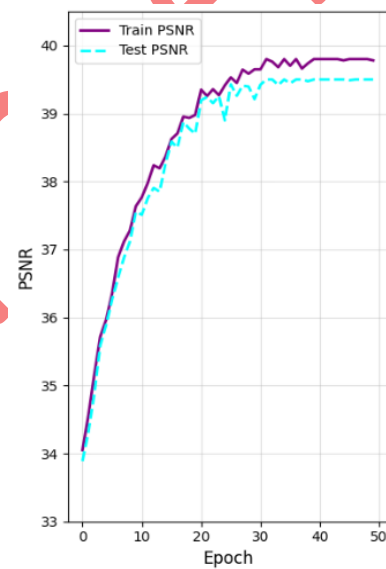
c

d

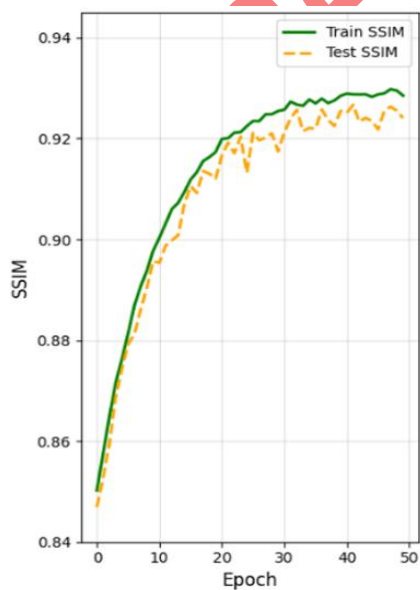
شکل ۱ شاخص های (a) جاکارد نزدیک به 0/875 که میانگین هر سه بخش میباشد، (b) نسبت سیگنال به نویز بالاتر از ۴۰ با یک عملکرد عالی، (c) شاخص شباهت ساختاری نزدیک به ۱ و نشاندهنده عملکرد عالی و (d) تابع هزینه روی داده های آزمایش مدل ترنسفورمر مرتبط با بازسازی تصویر



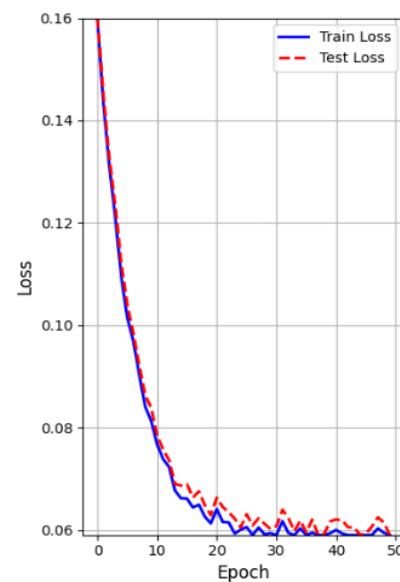
a



b



c



d

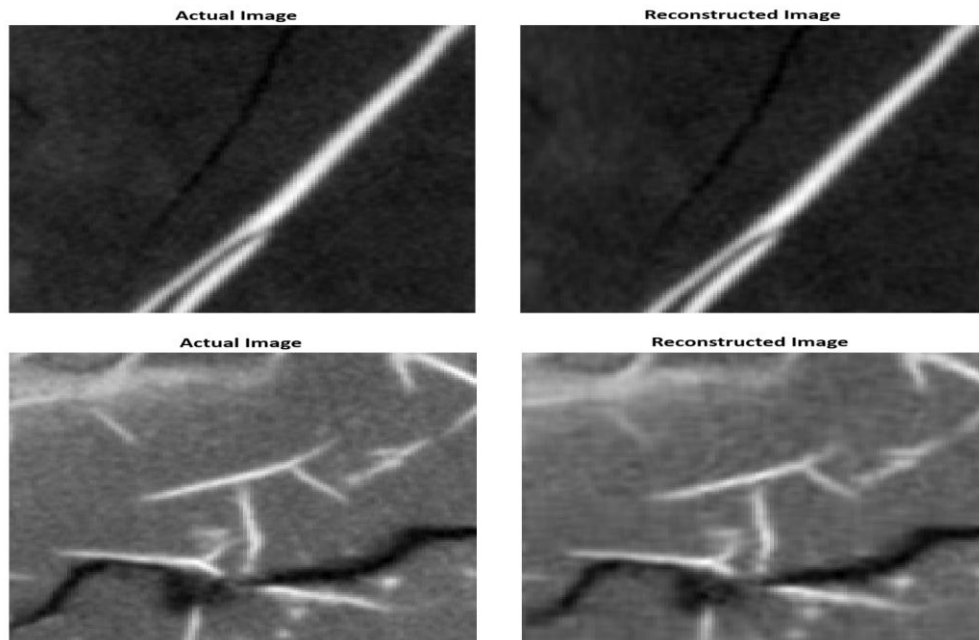
شکل ۲ شاخص های (a) جاکارد نزدیک به 80 که میانگین هر سه بخش میباشد، (b) نسبت سیگنال به نویز نزدیک به ۴۰ و نشاندهنده عملکرد خوب مدل، (c) شاخص شباهت ساختاری نزدیک به ۱ ولی مقداری کمتر از این شاخص در مدل ترنسفورمر بینایی و (d) تابع هزینه روی داده های آزمایش شبکه خودرمزگذار مرتبط با بازسازی تصویر

در ادامه، برای بررسی دقیق تر عملکرد دو معماری در مقیاس ریز و میزان حفظ میکروشکافها، یک ارزیابی دستی مبتنی بر مقیاس وکسل انجام شد که مبنای آن در ادامه بیان می شود. با توجه به ماهیت داده های استفاده شده، تصاویر روی یک شبکه ی گسسته از وکسل ها نمونه برداری می شوند؛ بنابراین کوچک ترین واحد مکانی قابل تفکیک در تصویر برابر با اندازه ی یک وکسل است. علاوه بر این، اثرات ذاتی تصویربرداری نظیر پخش شدگی نقطه ای، نویز و پدیده ی حجم جزئی باعث می شود تعیین دقیق موقعیت مرز شکستگی ها در تصویر مرجع با عدم قطعیتی در حدود ± 1 وکسل همراه باشد. از این رو، حفظ عرض شکستگی در بازسازی در محدوده ی خطای یک وکسل به عنوان معیار عملی و فیزیکی ارزیابی حفظ میکروشکافها در نظر گرفته شد.

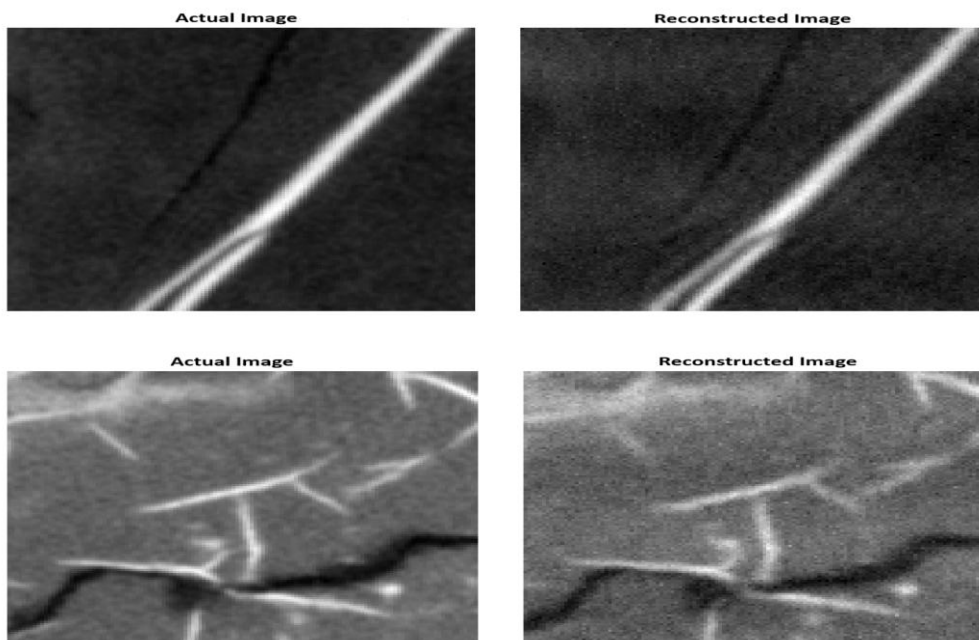
در معماری خودرمزگذار پیچشی، به دلیل ماهیت محلی کانولوشن ها و میدان دید موضعی فیلترها، ویژگی های ریزمقیاس نظیر لبه های شکستگی و بافت های نزدیک به مرز ماتریکس-شکاف به صورت مستقیم مدل سازی می شوند. ارزیابی دستی عرض میکروشکافها در نمونه های منتخب از طریق اندازه گیری پیکسل به پیکسل و تبدیل آن به مقیاس واقعی بر اساس اندازه ی وکسل دیتاست نشان داد که در بسیاری از موارد اختلاف عرض شکستگی بین تصویر مرجع و تصویر بازسازی شده ی خودرمزگذار در محدوده ی خطای یک وکسل باقی می ماند.

در مقابل، معماری ترنسفورمر بینایی به واسطه ی مکانیزم خودتوجه قادر است وابستگی های دوربرد بین تصاویر را بدون محدودیت میدان دید موضعی مدل کند؛ بنابراین پیوستگی طولی شکستگی ها، جهت گیری و اتصال ساختارهای چندمقیاسی را دقیق تر بازسازی می نماید. ارزیابی دستی عرض و پیوستگی میکروشکافها در نمونه های منتخب (با همان رویکرد پیکسل به پیکسل و تبدیل به مقیاس واقعی بر مبنای وکسل دیتاست) نشان داد که ترنسفورمر بینایی در اغلب موارد اختلاف عرض شکستگی را در حد خطای یک وکسل حفظ کرده و همزمان با کاهش گسست های مصنوعی، مسیرهای شکستگی را پیوسته تر و طبیعی تر بازسازی می کند.

شکل‌های مقایسه‌ای بین تصاویر واقعی و بازسازی‌شده برای هر دو مدل نیز مؤید همین نتایج بازسازی تصاویر هستند. در این تصاویر مشاهده می‌شود که معماری ترنسفورمر بینایی جزئیات ظریف‌تر شکستگی‌ها و اتصالات بین شکاف‌ها را با دقت بالاتری بازسازی کرده (شکل ۳) و نویز تصویری کمتری تولید نموده است. در حالی که مدل خودرمزگذار در تصاویر بازسازی‌شده نسبت به تصاویر اصلی، نواحی خاکستری دارای بافتی دانه‌دانه‌تر و نویز بیشتری هستند (شکل ۴).

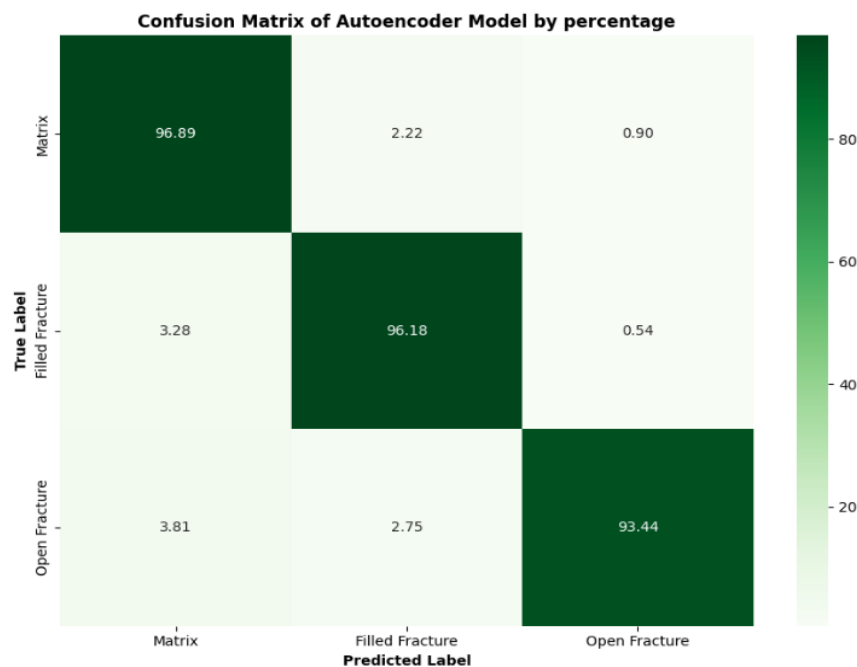


شکل ۳ تصاویر واقعی و بازتولید شده توسط مدل ترنسفورمر روی داده‌های اعتبارسنجی

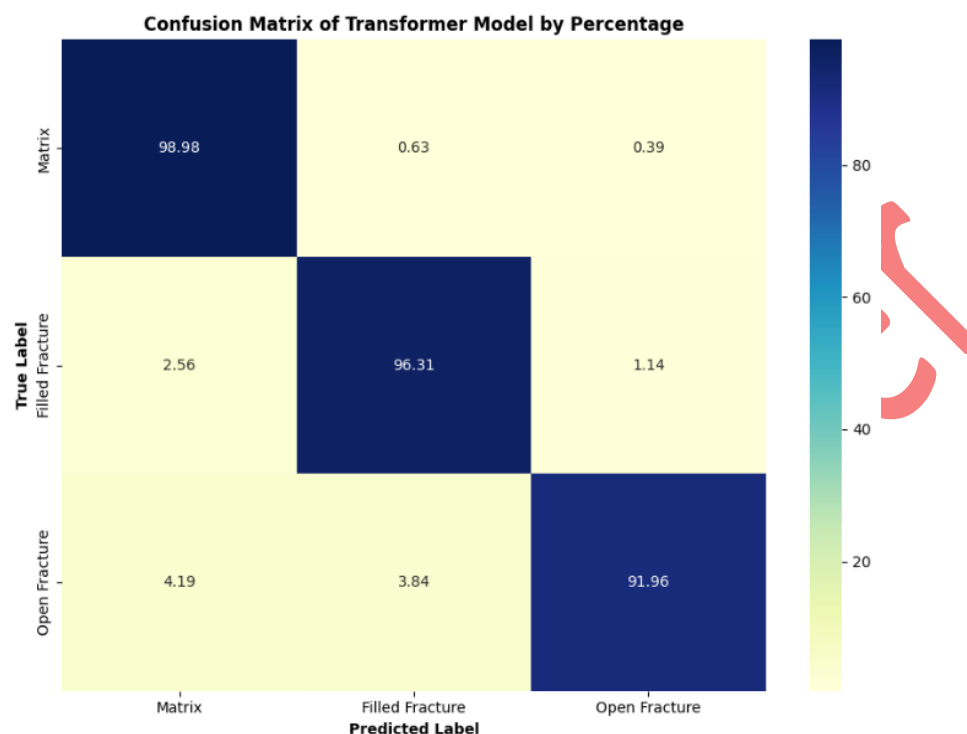


شکل ۴ تصاویر واقعی و بازتولید شده توسط شبکه خودرمزگذار روی داده‌های اعتبارسنجی

علاوه بر ارزیابی دیداری، ماتریس‌های سردرگمی برای سه کلاس تعریف شده (ماتریکس، شکاف باز و شکاف پر شده) نیز محاسبه شد. در مقایسه‌ی دو معماری خودرمزگذار و ترنسفورمر بینایی، هر دو در تشخیص سه دسته‌ی «ماتریکس»، «شکاف پر شده» و «شکاف باز» عملکرد بسیار خوبی دارند، اما تفاوت‌های ظریفی میان آن‌ها دیده می‌شود. معماری ترنسفورمر بینایی دقت کلی بالاتری دارد و در تشخیص نواحی مربوط به ماتریکس بسیار دقیق‌تر عمل کرده است (حدود ۹۹ درصد در برابر ۹۷ درصد در معماری خودرمزگذار). در مقابل، معماری خودرمزگذار در شناسایی شکستگی‌های باز کمی بهتر بوده است (۹۳/۴ درصد در برابر ۹۲ درصد). در مجموع، ترنسفورمر بینایی در جدا کردن بخش‌های ماتریکس از نواحی شکستگی‌ها عملکرد برتری دارد، اما خودرمزگذار در حفظ توازن بین همه‌ی دسته‌ها و در تفکیک نوع شکستگی‌ها (پر شده و باز) اندکی دقیق‌تر و پایدارتر عمل کرده است. در نتیجه، به طور کلی معماری ترنسفورمر بینایی نتیجه‌ی بهتری را نشان داده است (شکل ۵ و ۶).



شکل ۵ ماتریس‌های سردرگمی روی داده‌های آزمایش شبکه خودرمزنگار مرتبط با بخش بندی تصویر



شکل ۶ ماتریس های سردرگمی روی داده های آزمایش شبکه ترنسفورمر بینایی مرتبط با بخش بندی تصویر

در مقایسه‌ی دو معماری خودرمزگذار و ترنسفورمر بینایی، هر دو در تشخیص سه دسته‌ی «ماتریکس»، «شکستگی پرشده» و «شکستگی باز» عملکرد بسیار خوبی دارند، اما تفاوت‌های ظریفی میان آن‌ها دیده می‌شود. معماری ترنسفورمر بینایی در شاخص‌های دقت، یادآوری و امتیاز-F1 به‌ویژه در طبقه‌ی «ماتریکس»، عملکرد بهتری داشته است (به ترتیب ۰/۹۸۹، ۰/۹۹۰ و ۰/۹۸۹ در برابر ۰/۹۸۰، ۰/۹۶۹ و ۰/۹۷۴ در معماری خودرمزگذار) (جدول ۶). در طبقه‌ی «شکستگی پرشده» نیز ترنسفورمر بینایی با امتیاز-F1 برابر با ۰/۹۶۳ نسبت به خودرمزگذار با مقدار ۰/۹۶۱، عملکردی اندکی بهتر نشان داده است. در حالی که در طبقه‌ی «شکستگی باز» تفاوت میان دو مدل کمتر است و مقادیر امتیاز-F1 به ترتیب ۰/۹۲۱ برای ترنسفورمر بینایی و ۰/۹۳۷ برای خودرمزگذار به دست آمده است (جدول ۷). به‌طور کلی، معماری ترنسفورمر بینایی به دلیل دقت و یادآوری بالاتر در تشخیص نواحی «ماتریکس» و «شکستگی پرشده»، عملکرد کلی بهتری از خودرمزگذار دارد، هرچند خودرمزگذار در تفکیک دقیق‌تر نواحی «شکستگی باز» اندکی موفق‌تر بوده است.

جدول ۶ معیارهای گزارش طبقه بندی روی داده های آزمایش معماری ترنسفورمر بینایی مرتبط با بخش بندی تصویر

برچسب	Precision	Recall	F1-Score	شاخص جاکارد
(0) ماتریکس	0/989	0/990	0/989	0/92
(1) شکستگی پر شده	0/974	0/953	0/963	0/90
(2) شکستگی باز	0/923	0/920	0/921	0/85

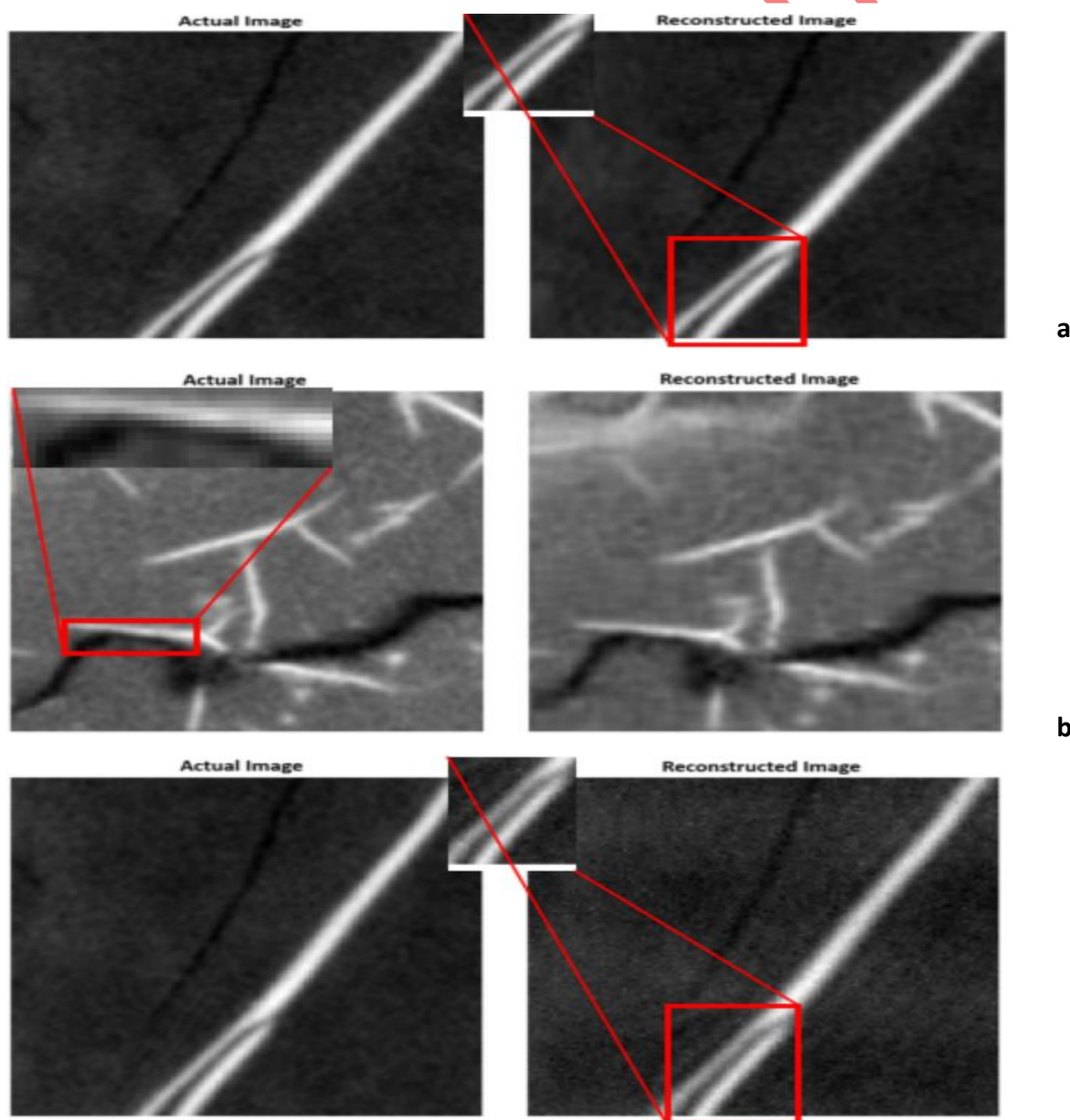
جدول ۷ معیارهای گزارش طبقه بندی روی داده های آزمایش معماری شبکه خودرمزگذار مرتبط با بخش بندی تصویر

برچسب	Precision	Recall	F1-Score	
(0) ماتریکس	0/980	0/969	0/974	0/83
(1) شکستگی پر شده	0/960	0/962	0/961	0/81
(2) شکستگی باز	0/940	0/934	0/937	0/80

از نظر تابع هزینه، منحنی های آموزش نشان دادند که معماری ترنسفورمر بینایی به پایداری همگرایی بیشتری رسیده و مقدار هزینه نهایی پایین تری نسبت به خودرمزگذار داشته است، که این امر نشانگر توانایی بالاتر در یادگیری روابط پیچیده بین تصاویر تصویر است. در افق آینده، به کارگیری چنین مدل هایی می تواند علاوه بر افزایش دقت بازسازی تصاویر، منجر به استخراج دقیق تر ویژگی های پتروفیزیکی همچون تخلخل دوگانه، نفوذپذیری مؤثر، جهت گیری شکستگی ها و پیوستگی شبکه ی شکاف ها شود [29]. در نتیجه، این رویکرد می تواند جایگزین روش های مخرب سنتی گردد، هزینه و زمان تحلیل را کاهش دهد و در نهایت بهره وری مدیریت مخازن پیچیده را بهبود بخشد. این پژوهش نشان می دهد که تلفیق یادگیری عمیق با داده های سیتی اسکن می تواند مسیر تحول در مدل سازی دیجیتال سنگ و درک رفتار دینامیکی مخازن شکاف دار را هموار کند.

نقاط قوت و محدودیت های مدل های بازسازی

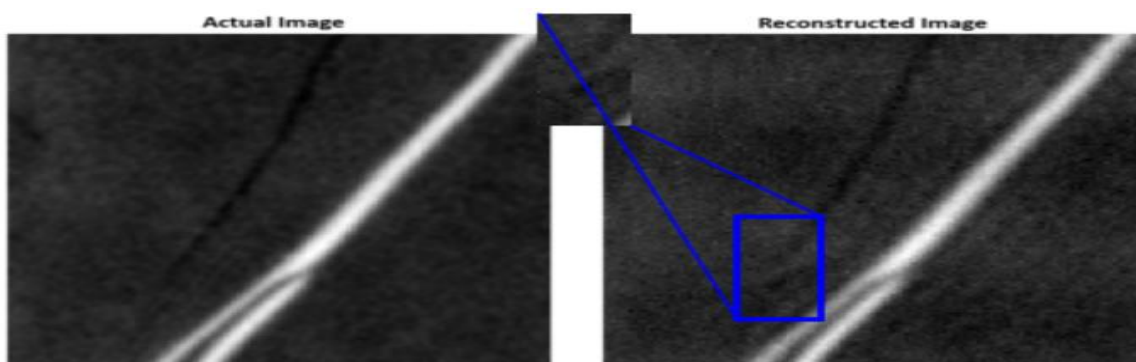
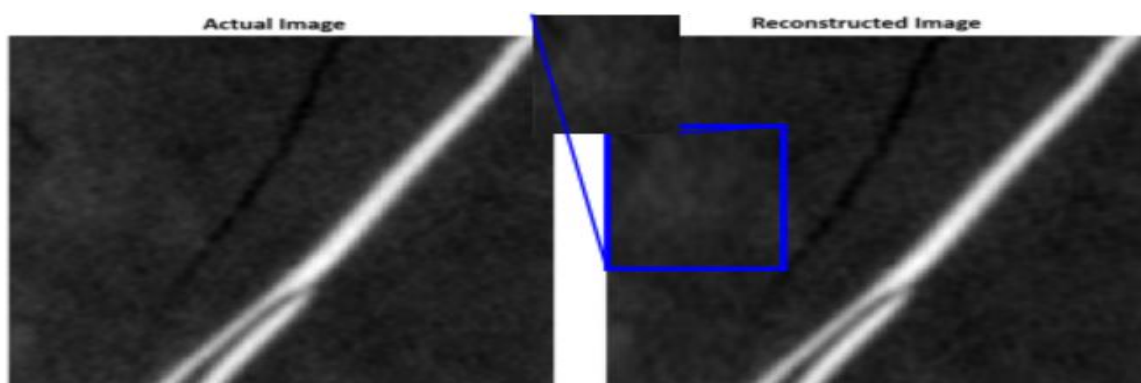
تصاویر شکل ۷ نمونه هایی از عملکرد دو مدل بازسازی در نواحی شکستگی را نشان می دهد. بزرگ نمایی های ارائه شده بیانگر آن است که مدل ترنسفورمر بینایی جزئیات شکستگی ها را با دقت بیشتری بازتولید می کند، در حالی که خودرمزگذار نیز در بازسازی الگوهای کلی ساختاری عملکردی پایدار و قابل قبول دارد. این مقایسه نشان می دهد هر دو مدل بسته به نوع ساختار سنگ، رفتار متفاوتی داشته و ویژگی های خاص خود را در فرآیند بازسازی ارائه می کنند.



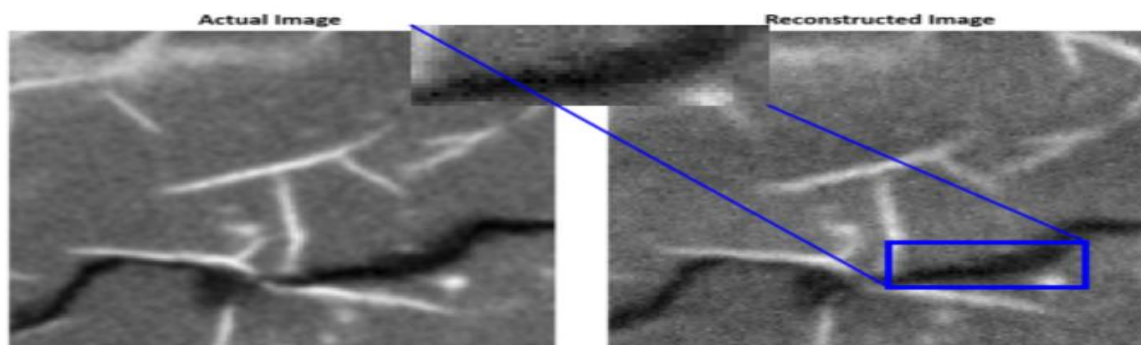
c

شکل ۷ a و b نشاندهنده نقاطی از تصویر که مدل ترنسفورمر بینایی توانسته به خوبی بازسازی انجام دهد و همچنین شکل c برای مدل خودرمزگذار که در بخش مورد نظر در این تصویر توانسته مانند مدل ترنسفورمر بینایی خوب عمل کند

همان گونه که در شکل ۸ مشاهده می شود، نواحی بزرگ نمایی شده نمونه هایی از بخش هایی هستند که بازسازی آن ها برای مدل ها چالش برانگیز بوده است. در تصویر a ، بخش مشخص شده در ناحیه ماتریکس همچنان مقداری نویز باقیمانده را نشان می دهد و مدل ترنسفورمر بینایی نتوانسته است این قسمت را به طور کامل صاف سازی و بازسازی کند. تصویر b نشان می دهد که مدل خودرمزگذار در ناحیه حاوی شکاف پر شده، بخشی از ساختار را به صورت محو بازسازی کرده و مرز شکستگی به طور



a



b

c

شکل ۸ a نشاندهنده نقاطی از تصویر که مدل ترنسفورمر بینایی نتوانسته به خوبی بازسازی انجام دهد (بخش ماتریکس مقداری نویز دارد) و شکل b تصویر بازسازی شده به کمک مدل خودرمزگذار و محدودیت بازسازی در قسمت مشخص. همچنین شکل c نشاندهنده محدودیت مدل خودرمزگذار با حذف مقداری از

کامل حفظ نشده است. همچنین در تصویر c، خروجی مدل خودرمزگذار نویز بیشتری را نشان می‌دهد و بخش علامت‌گذاری شده از شکاف خالی به‌طور دقیق بازتولید نشده است و مقداری نویز بیشتر در تصویر بازسازی شده به کمک خودرمزگذار وجود دارد.

این موارد بیانگر آن است که هر دو معماری، با وجود عملکرد کلی مناسب، در نواحی با هندسه پیچیده یا نسبت سیگنال به نویز پایین، با محدودیت‌هایی در بازسازی کامل ساختار مواجه هستند.

نقاط قوت و محدودیت‌ها

یکی از محدودیت‌های مهم این پژوهش به قابلیت تعمیم‌پذیری مدل‌ها مربوط می‌شود. داده‌های به‌کاررفته عمدتاً شامل تصاویر سیتی اسکن از نمونه‌های زغال‌سنگی شکاف‌دار در مقیاس ریز بوده‌اند؛ بنابراین تعمیم مستقیم نتایج به لیتولوژی‌های متفاوت نظیر سنگ‌های کربناته، ماسه‌سنگی یا شیل و نیز به داده‌هایی با اندازه و کسل، شدت تابش، اندازه متفاوت تصاویر یا شرایط تصویربرداری متفاوت، بدون ارزیابی مجدد قابل تضمین نیست. از آنجا که ویژگی‌های بافتی، تخلخل، هندسه شکستگی‌ها و الگوی نویز در هر نوع سنگ و هر مقیاس تصویربرداری می‌تواند به‌طور معناداری تغییر کند، در مواجهه با داده‌های جدید پدیده‌ی ناهمخوانی دامنه و مقیاس محتمل بوده و ممکن است عملکرد مدل را تحت تأثیر قرار دهد. با این حال، چون هر دو معماری خودرمزگذار و ترنسفورمر بینایی ذاتاً قابلیت یادگیری مقیاس‌ناوابسته دارند، در صورت فراهم بودن داده‌های هدف، می‌توان مدل‌های آموزش‌دیده را روی آن داده‌ها با تنظیم مجدد ابرپارامترها (Fine-Tuning) سازگار کرد. بر این اساس، ارزیابی تجربی روی مجموعه داده‌های متنوع‌تر، لیتولوژی‌های مختلف و مقیاس‌های تصویری بزرگ‌تر (از جمله داده‌های رخنمونی) به‌عنوان یک مسیر ضروری برای تثبیت پایداری و تعمیم‌پذیری روش در مطالعات آینده پیشنهاد می‌شود.

کارهای آینده

با توجه به عملکرد مطلوب مدل‌ها در بازسازی تصاویر مخازن شکاف‌دار، مسیرهای آینده می‌تواند بر کاربردپذیری نتایج در تحلیل‌های عملی مخزن متمرکز شود. نخست، استفاده از تصاویر بازسازی‌شده برای ساخت مدل‌های دیجیتالی سه‌بعدی دقیق‌تر شبکه شکستگی‌ها می‌تواند دقت شبیه‌سازی‌هایی مانند نفوذپذیری مؤثر، تخلخل دوگانه و جریان سیال را افزایش دهد. توسعه

یک مسیر خودکار شامل بازسازی، بخش‌بندی و استخراج ویژگی نیز امکان استفاده صنعتی از این روش را فراهم می‌کند. علاوه بر این، به‌کارگیری مدل‌های آموزش‌دیده در وظایفی مانند تشخیص شکستگی، ارزیابی آسیب مخزن و تحلیل عملیات‌هایی نظیر شکست هیدرولیکی از دیگر زمینه‌های توسعه است. ترکیب تصاویر با داده‌های مکمل نظیر برش نازک، NMR و FMI نیز می‌تواند به ایجاد یک مدل چندمنبعی دقیق‌تر برای توصیف مخزن منجر شود. در نهایت، تعمیم مدل به لیتولوژی‌ها، رزولوشن‌ها و شرایط نویری مختلف و همچنین توسعه نسخه‌های سبک‌تر برای استفاده در محیط‌های عملیاتی، از مهم‌ترین مسیرهای تحقیقاتی آینده به شمار می‌روند.

نتیجه‌گیری

در این پژوهش، با بهره‌گیری از دو معماری پیشرفته‌ی یادگیری عمیق شامل خودرمزگذار پیچشی و ترنسفورمر بینایی، فرآیند بازسازی و نویرزدایی تصاویر سی‌تی‌اسکن سنگ‌های مخازن شکاف‌دار طبیعی مورد بررسی قرار گرفت. هدف اصلی، بازتولید دقیق ساختارهای درونی شامل ماتریکس سنگ، شکاف‌های باز و شکاف‌های پرشده با مواد معدنی بود. نتایج اولیه نشان داد که هر دو مدل توانایی بازسازی ساختارهای شکاف‌دار را با دقت قابل‌قبولی دارند، با این حال ترنسفورمر بینایی در حفظ جزئیات ریز، مرزهای ساختاری و کاهش نویز، برتری محسوسی نسبت به خودرمزگذار داشت.

- از نظر کمی، ترنسفورمر بینایی در بازسازی تصاویر عملکرد برتری نسبت به خودرمزگذار پیچشی داشت. این مدل به نسبت سیگنال به نویز اوج ۴۵ و شاخص شباهت ساختاری ۰.۹۸ دست یافت، در حالی که خودرمزگذار به ترتیب مقادیر ۳۹ و ۰.۹۳ را ثبت کرد. بررسی شاخص جاکارد نیز نشان داد که ترنسفورمر در تفکیک دقیق‌تر نواحی تصویر به‌ویژه مرزبندی ماتریکس و شکاف‌های پرشده موفق‌تر بوده و توانسته است ارتباط‌های فضایی و ساختاری تصویر را بهتر حفظ کند، در حالی که مدل پیچشی در نواحی با جزئیات ریز دچار افت وضوح و از دست‌رفتن ساختار شده است.
- در تحلیل شاخص‌های بخش‌بندی سه‌کلاسه، ترنسفورمر بینایی برای کلاس ماتریکس مقادیر دقت، بازخوانی و امتیاز F1 را به ترتیب ۰.۹۸۹، ۰.۹۹۰ و ۰.۹۸۹/ به دست آورد. برای شکاف‌های پرشده، این مقادیر برابر با ۰.۹۷۴، ۰.۹۵۳ و ۰.۹۶۳۰ و برای شکاف‌های باز ۰.۹۲۳، ۰.۹۲۰ و ۰.۹۲۱ بود. در مقابل، خودرمزگذار پیچشی برای همین سه کلاس به ترتیب نتایج ۰.۹۸۰، ۰.۹۶۹ و ۰.۹۷۴، ۰.۹۶۰، ۰.۹۶۲ و ۰.۹۶۱، ۰.۹۴۰، ۰.۹۳۴ و ۰.۹۳۷ را ارائه داد. این مقایسه

- نشان می‌دهد که ترنسفورمر بینایی در شاخص‌های دقت و یادآوری به‌ویژه در طبقه ماتریکس و شکاف‌های پرشده عملکرد برتری دارد، هرچند خودرمزگذار در تشخیص برخی شکاف‌های باز باریک عملکرد نزدیک‌تری ارائه کرده است.
- از نظر کیفی، تصاویر بازسازی‌شده با ترنسفورمر بینایی دارای وضوح بیشتر، نویز کمتر و بازسازی دقیق‌تر شکستگی‌ها و اتصالات بودند. در مقابل، مدل خودرمزگذار خروجی‌هایی با بافت دانه‌دانه، خطوط محو و نویز بالاتر تولید کرد. ترنسفورمر توانست ساختارهای ظریف و شکستگی‌های ناپیوسته را با پیوستگی بهتری بازسازی کند که برای تحلیل دقیق شبکه شکاف‌ها بسیار مهم است.
 - بررسی منحنی‌های یادگیری دو مدل نشان داد که ترنسفورمر بینایی در طول فرایند آموزش از همگرایی پایدارتر و نرخ کاهش تابع هزینه‌ی سریع‌تری برخوردار است. این ویژگی بیانگر قابلیت بالای آن در یادگیری روابط پیچیده و وابستگی‌های بلندبرد بین تصاویر تصویر است، در حالی که خودرمزگذار در تکرارهای اولیه آموزش نوسانات بیشتری نشان داد و به حداقل محلی تابع هزینه نزدیک‌تر شد.
 - نتایج نشان دادند که علاوه بر نوع معماری، روش‌های پیش‌پردازش مانند آستانه‌گذاری چندسطحی اتسو و تقویت داده‌ها تأثیر زیادی در بهبود کیفیت بازسازی دارند. همچنین استفاده از ابرپارامترهای بهینه و آموزش روی GPU باعث افزایش دقت و سرعت همگرایی مدل‌ها شد، که برای پردازش داده‌های حجیم سیتی اسکن در تحلیل‌های پتروفیزیکی مانند تخلخل و نفوذپذیری اهمیت بالایی دارد.

مراجع

- [۱] Nguyen, K. L., Jumah, M., Verma, N., & Quttainah, R. (2023). Identifying Fracture Sweet Spots Using Artificial Neural Network Approach: A Case Study in Najmah/Sargelu Reservoir, Kra Al-Marū Field, West Kuwait. Middle East Oil, Gas and Geosciences Show, <https://doi.org/10.2118/213770-MS>
- [۲] Wilson, T. H., Smith, V., & Brown, A. (2015). Developing a model discrete fracture network, drilling, and enhanced oil recovery strategy in an unconventional naturally fractured reservoir using integrated field, image log, and three-dimensional seismic data. AAPG Bulletin, 99(4), 783–808. <https://doi.org/10.1306/10031414015>
- [۳] Sun, S., & Pollitt, D. A. (2021). Optimising development and production of naturally fractured reservoirs using a large empirical dataset. Petroleum Geoscience, 27(3), petgeo2020-079. <https://doi.org/10.1144/petgeo2020-079>
- [۴] Lyu, W., Zeng, L., Zhou, S., Du, X., Xia, D., Liu, G., Li, J., & Weng, J. (2019). Natural fractures in tight-oil sandstones: A case study of the Upper Triassic Yanchang Formation in the southwestern Ordos Basin, China. AAPG Bulletin, 103(1), 135–154. <https://doi.org/10.1306/0130191608617115>

- [٤] Hashemi, R., Saberi, F., Asoude, P., & Soleimani, B. (2024). Enhancing Reservoir Zonation through Triple Porosity System: A Case Study. *SPE Journal*, 29(06), 3043–3062. <https://doi.org/10.2118/219491-pa>
- [٥] Al-Rbeawi, S. (2020). The performance of complex-structure fractured reservoirs considering natural and induced matrix block size, shape, and distribution. *Journal of Natural Gas Science and Engineering*, 81, 103400. <https://doi.org/10.1016/j.jngse.2020.103400>
- [٦] Zhu, Y., Hou, D., Lan, M., Sun, X., & Ma, X. (2019). A comparison of ultra-high-resolution CT target scan versus conventional CT target reconstruction in the evaluation of ground-glass-nodule-like lung adenocarcinoma. *Quantitative Imaging in Medicine and Surgery*, 9(6), 1087–1094. <https://doi.org/10.2196/qims.2019.06.09>
- [٧] Faisal, R. S., Salih, N. M., Kamal, I., & Pr  at, A. (2023). X-ray computed tomography (CT) to scan the structure and characterize the mud cake incorporated with various magnetic NPs concentration: An application to evaluate the wellbore stability and formation damage. *Nanomaterials*, 13(12), 1843. <https://doi.org/10.3390/nano13121843>
- [٨] Cappuccio, F., Toy, V. G., Mills, S., & Adam, L. (2020). Three-dimensional separation and characterization of fractures in X-ray computed tomographic images of rocks. *Frontiers in Earth Science*, 8, 529263. <https://doi.org/10.3389/feart.2020.529263>
- [٩] Soltanmohammadi, R., & Faroughi, S. A. (2023). A comparative analysis of super-resolution techniques for enhancing micro-CT images of carbonate rocks. *Applied Computing and Geosciences*, 20, 100143. <https://doi.org/10.1016/j.acags.2023.100143>
- [١٠] Bizhani, M., Haeri Ardakani, O., & Little, E. (2022). Reconstructing high fidelity digital rock images using deep convolutional neural networks. *Scientific Reports*, 12, Article 4264. <https://doi.org/10.1038/s41598-022-08170-8>
- [١١] Shahmoradi, M., Rekabdar, G., & Adibi, S. (2016). A comparative study of new and current methods for dental micro-CT image denoising. *Dentomaxillofacial Radiology*, 45(3), 20150302. <https://doi.org/10.1259/dmfr.20150302>
- [١٢] Feng, X., Zhang, W., Su, X., & Xu, Z. (2021). Optical Remote Sensing Image Denoising and Super-Resolution Reconstructing Using Optimized Generative Network in Wavelet Transform Domain. *Remote. Sens.*, 13, 1858. <https://doi.org/10.3390/rs13091858>
- [١٣] Sahu, P., Kumar, A., & Rout, S. (2022). Wavelet guided 3D deep model to improve dental microfracture detection. In *Applied Medical Artificial Intelligence* (pp. 1–12). Springer. https://doi.org/10.1007/978-3-031-17721-7_16
- [١٤] Mathew, E. S., Singh, K., & Arbogast, B. (2025). Deep learning assisted denoising of fast polychromatic X-ray micro-CT imaging of multiphase flow in porous media. *Computers & Geosciences*, 204, 105990. <https://doi.org/10.1016/j.cageo.2025.105990>
- [١٥] Koetzier, L. R., Mastrodicasa, D., Szczykutowicz, T. P., van der Werf, N. R., Wang, A. S., Sandfort, V., van der Molen, A. J., Fleischmann, D., & Willemink, M. J. (2023). Deep Learning Image Reconstruction for CT: Technical Principles and Clinical Prospects. *Radiology*, 306(3), e221257. <https://doi.org/10.1148/radiol.221257>
- [١٦] Janssens, N., Huysmans, M., & Swennen, R. (2020). Computed Tomography 3D Super-Resolution with Generative Adversarial Neural Networks: Implications on Unsaturated and Two-Phase Fluid Flow. *Materials (Basel, Switzerland)*, 13(6), 1397. <https://doi.org/10.3390/ma13061397>
- [١٧] Shirmard, H., Farahbakhsh, E., Heidari, E., Beiranvand Pour, A., Pradhan, B., M  ller, D., & Chandra, R. (2022). A comparative study of convolutional neural networks and conventional machine learning models for lithological mapping using remote sensing data. *Remote Sensing*, <https://doi.org/10.3390/rs14040819>

- [١٩] Liu, L., Chang, B., Prodanović, M., & Pyrcz, M. J. (2025). AI-based digital rocks augmentation and assessment metrics. *Water Resources Research*, 61, e2024WR037939. <https://doi.org/10.1029/2024WR037939>
- [٢٠] Espinoza, D. N., Shovkun, I., Makni, O., & Lenoir, N. (2016). Natural and induced fractures in coal cores imaged through X-ray computed microtomography: Impact on desorption time. *International Journal of Coal Geology*, 154–155, 165–175. <https://doi.org/10.1016/j.coal.2015.12.012>
- [٢١] Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, 313(5786), 504–507. <https://doi.org/10.1126/science.1127647>
- [٢٢] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533–536. <https://doi.org/10.1038/323533a0>
- [٢٣] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2020). An image is worth 16×16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*. <https://arxiv.org/abs/2010.11929>
- [٢٤] Huynh-Thu, Q., & Ghanbari, M. (2008). Scope of validity of PSNR in image/video quality assessment. *Electronics Letters*, 44(13), 800–801. <https://doi.org/10.1049/el:20080522>
- [٢٥] Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4), 600–612. <https://doi.org/10.1109/TIP.2003.819861>
- [٢٦] Tatbul, N., Lee, T. J., Zdonik, S., Alam, M., & Gottschlich, J. (2018). Precision and recall for time series. In *Proceedings of the 32nd Conference on Neural Information Processing Systems (NeurIPS 2018)*, Montréal, Canada. <https://doi.org/10.48550/arXiv.1803.03639>
- [٢٧] Foody, G. M. (2023). Challenges in the real world use of classification accuracy metrics: From recall and precision to the Matthews correlation coefficient. *PLOS ONE*, 18(10), e0291908. <https://doi.org/10.1371/journal.pone.0291908>
- [٢٨] Wang, Z., Ning, X., & Blaschko, M. B. (2023). Jaccard metric losses: Optimizing the Jaccard index with soft labels. *Proceedings of the 37th Conference on Neural Information Processing Systems (NeurIPS 2023)*. <https://doi.org/10.48550/arXiv.2302.05666>
- [٢٩] Sobhani Oghaz, A., & Emami Niri, M. (2022). Effect of fracture geometrical properties on flow behavior and dynamic sensitivity analysis of fracture parameters in fractured reservoirs. *Petroleum Research*, 127, 49–64. <https://doi.org/10.22078/PR.2022.4893.3188>

A Comparative Study of Convolutional Autoencoder and Vision Transformer Architectures for Denoising and Reconstruction of Subsurface Fractured Reservoir Rock Images

Behrad Tabrizipour

Faculty of Petroleum Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

behrad.tabrizipour@srbiau.ac.ir

Abstract

Fractured reservoir rocks are among the most critical targets in petroleum engineering studies, as their complex networks of fractures and cracks play a decisive role in the porosity and permeability of these formations. CT-scan imaging is a key tool for analyzing the internal structure of such rocks; however, reduced image quality often makes accurate reconstruction challenging. In this study, two deep-learning architectures a Convolutional Autoencoder and a Vision Transformer were designed and compared with the objective of enhancing and reconstructing CT-scan images of fractured reservoir rocks. The primary task of the models was to reconstruct and denoise subsurface fractured images, and following reconstruction, a three-class segmentation process (matrix, open fracture, and filled fracture) was performed using the classical Otsu thresholding method to evaluate the improved performance of the reconstructed outputs. Quantitative and qualitative results demonstrated that the Vision Transformer outperformed the autoencoder due to its use of the attention mechanism. The Vision Transformer achieved a peak signal-to-noise ratio of 45 and a structural similarity index of 0.98 in image reconstruction, whereas the autoencoder achieved values of 39 and 0.93, respectively. In the Otsu-based three-class segmentation, the Vision Transformer again delivered superior performance in terms of precision, recall, and F1-score. For the matrix class, precision, recall, and F1-score reached 0.989, 0.990, and 0.989, compared to 0.980, 0.969, and 0.974 for the autoencoder. For the filled-fracture class, the results were 0.974, 0.953, and 0.963, and for the open-fracture class, 0.923, 0.920, and 0.921, while the autoencoder recorded 0.960, 0.962, 0.961 and 0.940, 0.934, 0.937 for the same classes. Overall, the findings indicate that the Vision Transformer is a more efficient and accurate option for CT-scan image reconstruction of fractured rocks, offering significant potential to improve numerical analyses of rock properties such as porosity and permeability in reservoir engineering studies.

Keywords: Fractured reservoir rock, image reconstruction, autoencoder, vision transformer, peak signal-to-noise ratio, structural similarity index

Accepted Paper