

ارائه یک شبکه مولد ترکیبی مبتنی بر توجه جهت افزایش وضوح تصاویر

محمد امین میرزایی^۱، محمدحسین شکور^۲

چکیده

بهبود کیفیت تصاویر در حوزه بینایی کامپیوتری به عنوان یکی از چالش‌های کلیدی مطرح شده است. در این مقاله، یک مدل جدید شبکه عصبی مولد متخاصم وضوح فوق‌العاده مبتنی بر توجه معرفی شده است که با ترکیب مدل‌های شبکه بهبود وضوح تصویر واقعی و شبکه توجه کانال‌محور رسوبی برای بهبود بازسازی تصاویر با وضوح بالا و کاهش اعوجاج‌ها و نویزهای پیچیده توسعه داده شده است. اعوجاج‌های تصویر شامل تحریف‌های هندسی، تاری و از بین رفتن جزئیات هستند که در تصاویر با وضوح پایین یا فشرده شده معمولاً مشاهده می‌شوند. در این مدل، قابلیت‌های شبکه بهبود وضوح تصویر واقعی برای کاهش نویز و افزایش وضوح به همراه توانایی شبکه توجه کانال‌محور رسوبی در حفظ جزئیات دقیق به کار گرفته شده‌اند. نتایج تجربی بر روی چند نمونه از مجموعه داده‌های شناخته شده در این حوزه نشان داده است که مدل پیشنهادی بهبود عملکردی حدود ۵ درصد به طور متوسط در معیار PIQE و ... را نسبت به روش‌های موجود مانند شبکه بهبود وضوح تصویر با تقویت توجه و شبکه بهبود وضوح تصویر واقعی در معیارهای شناخته شده بازسازی کور تصویر ارائه داده است. این بهبودها، مدل را قادر ساخته‌اند تا تصاویر را با وضوح بالا و حداقل اعوجاج و نویز بازسازی کند.

کلید واژه‌ها

بهبود کیفیت تصاویر، شبکه‌های مولد متخاصم، مکانیزم توجه کانالی، بازسازی تصاویر با وضوح بالا، اعوجاج تصویر، شبکه‌های ترکیبی، شبکه‌های عصبی عمیق

۱ - مقدمه

قدیمی و کم‌کیفیت، و همچنین تحلیل تصاویر در شرایط خاص، نقش قابل توجهی داشته است. در واقع، به دلیل کاربردهای فراوان این تکنیک در تسهیل فرآیندهای پزشکی و علمی، توجهات زیادی به خود جلب کرده است.

هدف اصلی از افزایش وضوح تک تصویری، بازسازی تصاویر با وضوح بالا از تصاویر با وضوح پایین است، به ویژه در شرایطی که تصویر دچار اعوجاج‌های پیچیده، ناشناخته یا نویزهای اضافی باشد. این فرآیند به گونه‌ای طراحی شده است که می‌تواند اطلاعات از دست رفته یا کم‌کیفیت را بازیابی کرده و تصویری واضح‌تر و دقیق‌تر ارائه دهد. به ویژه در شرایطی که تصاویر تحت تأثیر انواع مختلفی از اعوجاج‌ها و نویزها قرار گرفته‌اند، این روش‌ها قادر به بهبود وضوح و کیفیت تصاویر با استفاده از تکنیک‌های پیشرفته پردازش تصویر خواهند بود.

افزایش وضوح تصویر یکی از مسائل کلیدی و چالش برانگیز در حوزه بینایی کامپیوتری به شمار می‌رود که اهمیت ویژه‌ای در بسیاری از زمینه‌ها داشته است. این تکنیک به طور خاص جهت تولید تصاویر با وضوح بالا از تصاویر با وضوح پایین استفاده شده است و در حوزه‌های مختلفی مانند بهبود تصاویر، بازسازی تصاویر پزشکی، ارتقای کیفیت تصاویر ماهواره‌ای، بازسازی

این مقاله در آبان‌ماه ۱۴۰۳ دریافت شد؛ در بهمن‌ماه بازنگری و سپس پذیرفته گردید.

^۱ دانشجوی کارشناسی ارشد هوش مصنوعی و رباتیک، دانشکده فنی و مهندسی، گروه مهندسی کامپیوتر، دانشگاه اراک
رایانامه: m.mirzaei.01@msc.araku.ac.ir

^۲ استادیار، دانشکده فنی و مهندسی، گروه مهندسی کامپیوتر، دانشگاه اراک
رایانامه: mh-shakoor@araku.ac.ir

نویسنده مسئول: محمد امین میرزایی

در این پژوهش، با ترکیب ویژگی‌های کلیدی شبکه بهبود وضوح تصویر واقعی و شبکه توجه کانال‌محور رسوبی، مدلی ارائه شده است که توانایی بازسازی تصاویر با وضوح بالا و کیفیت بصری بهتری را داشته است.

این مدل از معماری بهینه‌شده شبکه بهبود وضوح تصویر واقعی برای حذف نویز و پردازش دقیق جزئیات بهره می‌برد، همچنین، مدل پیشنهادی با بهره‌گیری از بلوک توجه کانالی موجود در شبکه توجه کانال‌محور رسوبی، به استخراج و تمرکز بر کانال‌های مهم تصویری می‌پردازد و بدین ترتیب جزئیات ساختاری را با دقت بیشتری بازسازی می‌کند.

این ترکیب باعث شده است که مدل پیشنهادی در حفظ ساختارهای اصلی و افزایش وضوح تصاویر عملکردی عالی داشته باشد و بتواند به طور خاص در شرایط پیچیده و نویزی نتایج بسیار خوبی را ارائه دهد [۲, ۵, ۶].

این مدل جدید نه تنها توانسته است در معیارهای BRISQUE، PIQE، NIQE، بی نظمی و گرادیان که از معروف ترین معیارها در زمینه بازسازی تصاویر کور هستند به نتایج قابل توجهی دست یابد، بلکه در بازسازی تصاویر با وضوح بالا و حفظ جزئیات پیچیده نیز عملکردی فراتر از مدل‌های قبلی نشان داده است. بهبود کیفیت تصاویر در بسیاری از کاربردهای عملی، از جمله تصاویر پزشکی، تصاویر ماهواره‌ای و تصاویر نظارتی، می‌تواند به تحلیل‌های دقیق‌تر و تشخیص‌های بهتری منجر شود. مدل پیشنهادی به عنوان ابزاری مؤثر در بازسازی تصاویر با کیفیت بصری بالا و ساختارهای دقیق ارائه شده است [۷].

۲- مروری بر کارهای گذشته

در سال‌های اخیر، تقویت کیفیت تصاویر به عنوان یکی از موضوعات کلیدی در حوزه پردازش تصویر و بینایی ماشین مورد توجه قرار گرفته است. هدف اصلی این فرآیند، تبدیل تصاویر با وضوح پایین به تصاویر با وضوح بالا است، به گونه‌ای که جزئیات بیشتری از تصویر اصلی بازسازی شود و کیفیت بصری تصویر بهبود یابد [۸].

روش‌های مختلفی در این زمینه توسعه یافته‌اند که عمدتاً به دو دسته کلی تقسیم می‌شوند: روش‌های مبتنی بر بازسازی و روش‌های مبتنی بر یادگیری عمیق. در سال‌های اخیر، با پیشرفت روش‌های یادگیری عمیق، مدل‌های مبتنی بر شبکه‌های عصبی عمیق به ویژه شبکه‌های رسوبی و شبکه‌های مولد تخصصی به شدت مورد توجه قرار گرفته‌اند [۹].

اعوجاج‌های ناشی از محدودیت‌های سخت‌افزاری، نویزهای دیجیتال، یا تداخلات محیطی می‌تواند بر دقت و وضوح تصویر تأثیر منفی بگذارد؛ اما با استفاده از روش‌های افزایش وضوح تک تصویری، این مشکلات کاهش یافته و تصاویر دقیق‌تر و واضح‌تر تولید شده که به عنوان بازسازی کور تصویر شناخته می‌شوند [۱].

در طی سال‌های اخیر، مدل‌های متعددی برای حل این مسئله توسعه یافته‌اند که هر کدام به طور خاص به بهبود کیفیت بازسازی و افزایش وضوح تصویر پرداخته‌اند. یکی از مدل‌های برجسته در زمینه افزایش وضوح تک تصویری، شبکه بهبود وضوح تصویر واقعی [۲] است که برای بازسازی تصاویر با وضوح بالا در شرایط واقعی و نویزی طراحی شده است.

این مدل، نسخه‌ای بهینه‌شده از شبکه عصبی بهبود وضوح تصویر تقویت شده محسوب می‌شود که با به کارگیری بلوک‌های متراکم رسوبی تو در تو و با حذف نویزهای ناخواسته، توانسته است کیفیت تصاویر بازسازی شده را به طور چشمگیری افزایش دهد. این مدل از یک تابع هزینه ادراکی بهره می‌برد که باعث می‌شود جزئیات تصویر به طور طبیعی و واقعی بازسازی شوند و کیفیت بصری بهتری حاصل شود. این ویژگی‌ها موجب شده‌اند که شبکه بهبود وضوح تصویر واقعی در شرایط نویزی و کیفیت پایین عملکرد مناسبی داشته باشد و به یکی از انتخاب‌های اصلی در پردازش و بازسازی تصاویر دنیای واقعی تبدیل شود [۲, ۳].

شبکه توجه کانال‌محور رسوبی نیز یکی دیگر از مدل‌های پیشرفته در زمینه افزایش وضوح تک تصویری است که با بهره‌گیری از مکانیزم توجه کانالی، توانسته است به بهبود چشمگیری در بازسازی جزئیات تصویر دست یابد. این مدل از بلوک‌های توجه کانالی استفاده می‌کند که باعث می‌شود هر کانال تصویر به طور جداگانه مورد توجه قرار گیرد و اطلاعات کلیدی آن استخراج شود. این توجه کانالی به شبکه اجازه می‌دهد تا در بازسازی جزئیات پیچیده و حفظ ساختارهای اصلی تصویر دقت بیشتری داشته باشد و در عین حال نویزها و جزئیات نامطلوب را کاهش دهد.

این ویژگی شبکه توجه کانال‌محور رسوبی را به یکی از انتخاب‌های برتر برای بازسازی تصاویر با کیفیت بصری بالا تبدیل کرده است که می‌تواند جزئیات تصویر را به طور دقیق‌تر بازسازی کند و تصاویر بازسازی شده از وضوح و کیفیت بیشتری برخوردار شوند [۴, ۵].

Blind Image Super Resolution

Real Enhanced Super Resolution Generative Adversarial Network (Real ESRGAN)

Enhanced Super-Resolution Generative Adversarial Network (ESRGAN)

Residual in Residual Dense Block (RRDB)

Perceptual Loss

Residual Channel Attention Network (RCAN)

Naturalness Image Quality Evaluator

Perception-based Image Quality Evaluator

Blind Referenceless Image Spatial Quality Evaluator

Residual Networks

با الگوهای تکراری و ساختارهای پیچیده دارد. این ویژگی‌ها موجب شد که این مدل به یکی از مدل‌های برجسته در بازسازی تصاویر تبدیل شود و توجه محققان را به خود جلب کند [۱۷، ۱۸].

در سال ۲۰۱۹، مدل DRLN^۵ به‌عنوان یک پیشرفت چشمگیر در زمینه تقویت کیفیت تصاویر معرفی شد. این مدل از ساختار لایه‌ای بازگشتی عمیق برای بهبود دقت و بازسازی تصاویر بهره می‌برد. این مدل با استفاده از لایه‌های بازگشتی عمیق، می‌تواند ویژگی‌های مهم تصویر را استخراج و بافت‌های پیچیده را بازسازی کند.

این مدل به دلیل ساختار بازگشتی خود قادر است تا با تعداد پارامترهای کمتر، اطلاعات بیشتری از تصاویر ورودی را بازسازی کند و به جزئیات دقیق‌تری دست یابد. ویژگی برجسته این مدل، استفاده از لایه‌های بازگشتی است که باعث می‌شود مدل، اطلاعات بیشتری را از لایه‌های قبلی به لایه‌های بعدی منتقل کند و در نتیجه، جزئیات بیشتری را در بازسازی تصویر حفظ کند. این قابلیت، این مدل را به یکی از مدل‌های مؤثر در پردازش تصاویر پیچیده و بافت‌های گوناگون تبدیل کرده است [۱۹، ۲۰].

در همان سال، شبکه عصبی توجه کانال‌محور رسوبی توسط ژانگ و همکارانش معرفی شد که یکی از موفق‌ترین مدل‌ها در تقویت کیفیت تصاویر محسوب می‌شود. این مدل با بهره‌گیری از مکانیزم توجه کانالی^۶ توانست به استخراج دقیق‌تر و بهتر ویژگی‌های تصویر بپردازد. مکانیزم توجه کانالی به این مدل اجازه می‌دهد که در فرآیند بازسازی، تمرکز بیشتری بر کانال‌های تصویری داشته باشد که جزئیات بیشتری دارند و در نتیجه، دقت بازسازی افزایش یابد. این ساختار، این مدل را به یکی از مدل‌های پیشرفته و مؤثر برای بازسازی جزئیات دقیق و افزایش وضوح تصاویر تبدیل کرده است [۴، ۵].

پس از موفقیت این مدل‌ها، شبکه‌های مولد تخصصی به عنوان یک رویکرد قوی برای تولید تصاویر با کیفیت بالا معرفی شدند. شبکه‌های مولد تخصصی که در سال ۲۰۱۴ توسط گوگولوو و همکارانش ارائه شدند، به سرعت به یکی از روش‌های مؤثر در تولید تصاویر و به خصوص تقویت کیفیت تصاویر تبدیل شده است. این شبکه‌ها از دو بخش اصلی تشکیل شده‌اند: مولد^۷ گه تصاویر جدیدی تولید می‌کند و تفکیک کننده^۸ گه سعی می‌کند تفاوت بین تصاویر واقعی و تصاویر تولید شده را تشخیص دهد. رقابت بین این دو بخش منجر به ایجاد شبکه‌هایی می‌شود که توانایی تولید تصاویر با کیفیت بالا را دارند. این روش، زمینه‌ساز

در این میان، مدل‌های مبتنی بر شبکه‌های مولد تخصصی به دلیل ساختار و معماری منحصر به فرد خود در تولید تصاویر با کیفیت بالا پیشرو شده‌اند [۱۰].

روش تقریب بهترین صفحه که با هدف تخمین پیکسل‌های مورد نیاز در بزرگ‌نمایی تصاویر با استفاده از تقریب بهترین صفحه و سازگاری با بافت تصویر توسط علی محمد اسمعیلی زینی در سال ۲۰۱۷ معرفی گردیده است [۱۱].

مدل شبکه عصبی عمیق سلسله مراتبی خودتوجه و رمزگذاری پویای تک متغیره که با هدف بهبود وضوح تصاویر و بهینه‌سازی معماری شبکه‌های بزرگ‌نمایی تصاویر با استفاده از بلوک‌های سلسله‌مراتبی و توجه خودآموز توسط علی دلشادی و همکارانش در سال ۲۰۲۳ معرفی گردیده است [۱۲].

در سال ۲۰۱۴ شبکه عصبی کانولوشنی برای بازسازی تصویر SRCNN^۹ توسط دونگ و همکاران، به‌عنوان نخستین مدل مبتنی بر شبکه‌های عصبی عمیق برای بازسازی تصویر معرفی شد. این مدل با استفاده از نگاشت غیرخطی و کانولوشن مستقیم، کیفیت بازسازی بهتری نسبت به روش‌های کلاسیک ارائه داد [۱۳].

در سال ۲۰۱۶، مدل شبکه عصبی VDSR^{۱۰} معرفی شد که با به‌کارگیری لایه‌های کانولوشنی متعدد و بهبود سرعت همگرایی، دقت بالاتری در بازسازی تصویر فراهم کرد [۱۴].

یکی از نخستین مدل‌های موفق در این حوزه، EDSR^{۱۱} است که در سال ۲۰۱۷ توسط لیم و همکارانش معرفی شد. این مدل از معماری عمیق و ساده‌تری نسبت به مدل‌های پیچیده‌ی قبلی بهره برده است و با حذف لایه‌های اضافی و تمرکز بر استفاده بهینه از بلوک‌های رسوبی، توانسته است به‌طور چشمگیری کیفیت بازسازی تصاویر را بهبود بخشد.

این مدل موفق شده است که در رقابت‌های بین‌المللی مرتبط با تقویت کیفیت تصاویر، نتایج بسیار مطلوبی کسب کند و به عنوان یکی از مرجع‌های اصلی در این زمینه شناخته شده است. ویژگی شاخص این مدل، در سادگی ساختار و توانایی بالای آن در بازسازی دقیق بافت‌ها و جزئیات پیچیده در تصاویر بوده است [۱۵، ۱۶].

در ادامه‌ی توسعه این روش‌ها، مدل MSRN^{۱۲} در سال ۲۰۱۸ معرفی شد. این مدل برای بهبود وضوح تصاویر، از بلوک‌های رسوبی چندمقیاسی بهره می‌برد که به شبکه این امکان را می‌دهد تا اطلاعات تصویر را در مقیاس‌های مختلف پردازش کند. این مدل قادر است جزئیات دقیق‌تر و پیچیده‌تری را از تصویر استخراج کرده و به همین دلیل عملکرد بهتری در بازسازی تصاویر

^۵Deep Recursive Layer Network

^۶Channel Attention Mechanism

^۷Generator

^۸Discriminator

^۹Super-Resolution Convolutional Neural Network

^{۱۰}Very Deep Super-Resolution Network

^{۱۱}Enhanced Deep Super-Resolution

^{۱۲}Multi-Scale Residual Network

در مواقعی که اطلاعات اولیه محدود است. همچنین، مدل‌های فعلی در حفظ ساختارهای پیچیده و الگوهای تکراری در تصاویر با محدودیت‌هایی مواجه هستند.

به همین دلیل، توسعه مدل‌های جدید که بتوانند این چالش‌ها را برطرف کنند، اهمیت فراوانی دارد. مدل‌های جدیدی که در حال حاضر در دست توسعه‌اند، از تکنیک‌های مختلفی مانند شبکه‌های عصبی عمیق‌تر، مکانیزم‌های توجه پیچیده‌تر استفاده می‌کنند تا بتوانند کیفیت تصاویر تولید شده را بهبود دهند. هدف این تحقیقات، ارائه مدل‌هایی است که نه تنها تصاویر با وضوح بالا تولید کنند، بلکه جزئیات دقیق و کیفیت بصری بهتری نیز داشته باشند [۱۰].

۳- روش پیشنهادی

مدل پیشنهادی در این مقاله با تلفیق نقاط قوت دو مدل شناخته شده شبکه توجه کانال‌محور رسوبی شبکه بهبود وضوح تصویر واقعیهیوهای قابل توجهی را به همراه داشته است. ایده اصلی ترکیب این دو معماری بر مبنای استفاده از قابلیت‌های تکمیلی آن‌ها در بهبود کیفیت تصاویر ارائه شده است.

شبکه عصبی توجه کانال‌محور رسوبیبه دلیل ساختار بلوک‌های توجه کانالی خود، در حفظ جزئیات و دقت بازسازی تصاویر عملکرد برتری داشته است، در حالی که شبکه بهبود وضوح تصویر واقعیهیینه‌سازی‌های موثری را برای کاهش نویز و اعوجاج داشته است. ترکیب این دو مدل به منظور بهره‌گیری همزمان از این مزایا، راهکاری برای تولید تصاویر با وضوح بالا و حداقل اعوجاج فراهم کرده است.

ایده تلفیق این دو معماری از مشاهده نقاط ضعف و قوت مدل‌های موجود در حوزه بازسازی تصویر و تحلیل چالش‌های باقی‌مانده ناشی شده است. بررسی عملکرد شبکه بهبود وضوح تصویر واقعیهیانشان داده است که اگرچه این مدل به دلیل استفاده از تابع تخریب‌آو تمرکز بر کاهش نویز و تاری، در حذف این مشکلات بسیار کارآمد است، اما برای طبیعی‌تر کردن تصاویر، توجه کمتری به حفظ جزئیات پیچیده دارد و جزئیات کمتری را بازسازی کرده است.

از سوی دیگر، شبکه توجه کانال‌محور رسوبیبا استفاده از بلوک توجه کانالی، تمرکز بیشتری بر شناسایی و تقویت ویژگی‌های مهم تصویر داشته و در حفظ دقت جزئیات عملکرد بهتری نشان داده است، اما در کاهش نویز بهینه عمل نکرده است. بنابراین، ایده استفاده از بلوک‌های ترکیبی با الهام از ساختارهای کلیدی هر دو معماری مطرح شده و هدف از آن بهبود همزمان هر دو جنبه کیفیت تصویر بوده است.

توسعه مدل‌های جدیدی در تقویت کیفیت تصاویر شد که توانستند به بازسازی جزئیات دقیق‌تری دست یابند [۲۱, ۲۲].

یکی از مدل‌های پیشگام در این زمینه، شبکه عصبی بهبود وضوح تصویر تقویت شده بود که در سال ۲۰۱۸ معرفی شد. این مدل به‌طور خاص برای بازسازی تصاویر با کیفیت بالا از طریق معماری شبکه مولد تخصصی طراحی شد. این مدل توانست تصاویری با کیفیت بصری قابل توجه و وضوح بالایی را تولید کند. این مدل از بلوک‌های متراکم رسوبی تو در تو استفاده می‌کند که موجب بهبود دقت و پایداری شبکه در بازسازی تصاویر می‌شود.

علاوه بر این، استفاده از توابع هزینه مانند تابع هزینه ادراکی که شباهت‌های ادراکی بین تصاویر بازسازی‌شده و تصاویر واقعی را اندازه‌گیری می‌کند، به بهبود کیفیت بصری تصاویر کمک شایانی کرده است. این ویژگی‌ها موجب شد که این مدل به‌عنوان یک مدل مرجع در این حوزه شناخته شود [۲۳].

پس از موفقیت این مدل، شبکه عصبی بهبود وضوح تصویر واقعی در سال ۲۰۲۱ معرفی شد که نسخه‌ای پیشرفته و کاربردی‌تر از شبکه عصبی بهبود وضوح تصویر تقویت شده است. این مدل به‌طور ویژه برای بازسازی تصاویر در شرایط واقعی و با کیفیت پایین طراحی شده است. این مدل از ساختارهای پیچیده‌تر توجه و از بلوک‌های رسوبی بیشتری استفاده می‌کند تا بازسازی دقیق‌تری از تصاویر با نویز و کیفیت پایین انجام دهد. شبکه بهبود وضوح تصویر واقعی می‌تواند تصاویر با کیفیت پایین را حتی در شرایط نویزی بازسازی کرده و جزئیات دقیق‌تری را ارائه دهد. این ویژگی‌ها این مدل را به یک ابزار قدرتمند برای تقویت کیفیت تصاویر در شرایط واقعی تبدیل کرده است [۲, ۳].

در سال ۲۰۲۳، شبکه عصبی بهبود وضوح تصویر با تقویت توجه^۹ به‌عنوان یک پیشرفت جدید در حوزه بازسازی تصاویر با وضوح پایین معرفی شد. این مدل با بهره‌گیری از ساختار U-Net مبتنی بر توجه و یک شبکه مولد متخاصم^{۱۰} چند مقیاسی، به بهبود کیفیت تصاویر و بازسازی دقیق‌تر جزئیات کمک می‌کند.

این مدل به‌طور خاص برای مواجهه با چالش‌های موجود در تصاویر دچار اعوجاج طراحی شده و با استفاده از تکنیک‌های پیشرفته، قادر به شناسایی و حفظ ویژگی‌های ساختاری تصاویر در مقیاس‌های مختلف است. این مدل به‌واسطه انطباق با نیازهای خاص هر تصویر، کیفیت بصری را به‌طور قابل ملاحظه‌ای افزایش می‌دهد و به‌عنوان یکی از پیشرفته‌ترین رویکردها در زمینه تقویت کیفیت تصاویر شناخته می‌شود [۶].

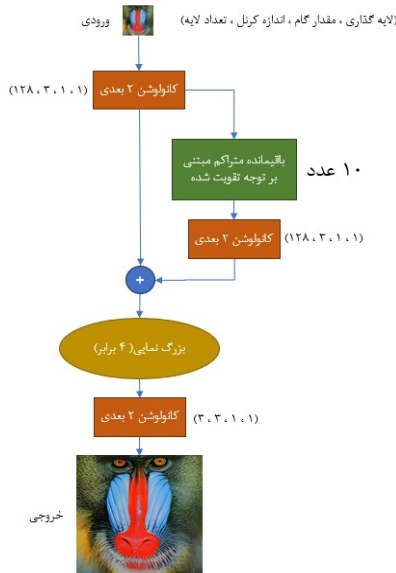
با وجود موفقیت‌های چشمگیر این مدل‌ها، همچنان چالش‌هایی در زمینه تقویت کیفیت تصاویر وجود دارد. یکی از چالش‌های اصلی، بازسازی دقیق و واقعی جزئیات است، به‌ویژه

^۹Attention based Enhanced Super-Resolution Generative Adversarial Network
^{۱۰}Generative Adversarial Network

۳-۱ ساختار مولد

ساختار مولد مدل پیشنهادی از اجزای متعددی تشکیل شده است که هر کدام نقشی مهم در فرآیند بازسازی تصاویر ایفا می‌کنند. نوآوری اصلی این ساختار در استفاده از ترکیب بلوک‌های توجه کانالی و بلوک‌های متراکم رسوبی که به منظور تقویت ویژگی‌های مهم تصویر و جلوگیری از ناپدید شدن گرادیان‌ها طراحی شده‌اند.

در این طراحی، از قدرت حفظ جزئیات در معماری شبکه توجه کانال محور رسوبی و توانایی کاهش نویز شبکه بهبود وضوح تصویر واقعی بهره‌برداری شده است. شکل ۲ ساختار مولد پیشنهادی را نشان می‌دهد.



شکل ۲: ساختار مولد در مدل پیشنهادی

۳-۱-۱ بلوک بزرگ‌نمایی

بلوک بزرگ‌نمایی: به منظور افزایش وضوح تصویر طراحی شده است. این بلوک شامل دو مرحله اصلی است [۲۳]:

- **افزایش ابعاد تصویر:** در این مرحله، با استفاده از یک لایه بزرگ‌نمایی، ابعاد تصویر به میزان دلخواه (معمولاً دو برابر) افزایش می‌یابد. این لایه با استفاده از روش نزدیکترین همسایگی عملیات بزرگ‌نمایی را انجام می‌دهد.
- **لایه کانولوشن و تابع فعال‌سازی:** پس از افزایش ابعاد، یک لایه کانولوشنی با هسته 3×3 و یک تابع فعال‌سازی LeakyReLU اعمال می‌شود که به تقویت جزئیات تصویر و بهبود کیفیت آن کمک می‌کند. این فرآیند به شبکه اجازه می‌دهد تا تصاویر ورودی با وضوح پایین را به تصاویر خروجی با وضوح بالا تبدیل کند.

جهت رسیدن به این هدف، تغییرات زیر به صورت یکپارچه و بهینه‌سازی شده به کار گرفته شده‌اند:

۱. تغییرات معماری مولد:

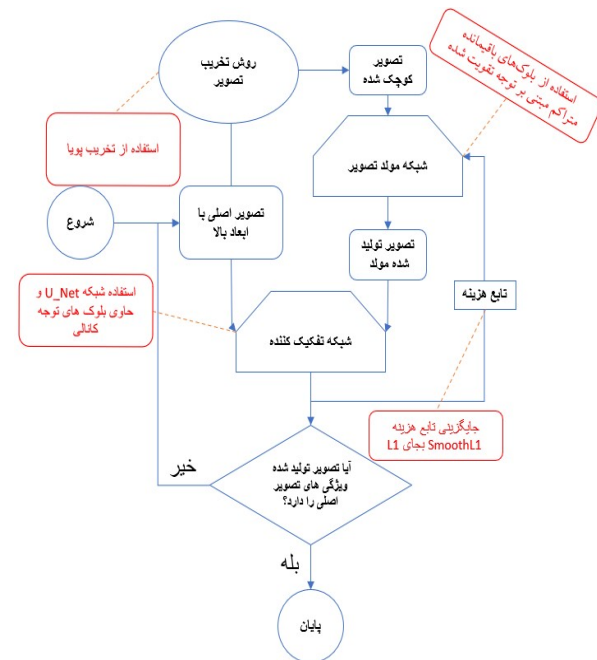
افزافه شدن بلوک‌های رسوبی متراکم مبتنی بر توجه تقویت شده که تلفیقی از بلوک‌های رسوبی توجه کانالی در شبکه توجه کانال محور رسوبی بلوک‌های متراکم رسوبی تو در تو در شبکه بهبود وضوح تصویر واقعی هستند. این بلوک‌ها با هدف تقویت ویژگی‌های مهم و جلوگیری از ناپدید شدن گرادیان‌ها در طول آموزش طراحی شده‌اند.

۲. تغییرات در تابع هزینه:

ترکیب توابع هزینه مختلف برای بهبود کیفیت بازسازی، شامل تابع هزینه مبتنی بر ادراک، تابع هزینه Smooth L1 برای کاهش خطاهای بزرگ و کوچک به صورت بهینه.

۳. استفاده از شبکه تفکیک‌کننده بهبودیافته:

طراحی شبکه تفکیک‌کننده مبتنی بر U-Net با بلوک‌های توجه کانالی و کانولوشن، برای افزایش دقت در تشخیص تفاوت بین تصاویر واقعی و تولیدشده و بهبود عملکرد مولد.



شکل ۱: دیاگرام کلی روش پیشنهادی و تغییرات انجام شده جهت رسیدن به مدل پیشنهادی

در شکل ۱ دیاگرام کلی روش پیشنهادی و تغییرات انجام شده و ارتباط هر قسمت از مدل را نشان داده شده است که در ادامه توضیحات کامل هر قسمت ارائه شده است.

۳-۱-۲ تابع هزینه مولد

تابع هزینه در مدل مولد پیشنهادی، مشابه شبکه بهبود وضوح تصویر واقعی طراحی شده و شامل بخش‌های مختلفی است که هر کدام نقش مهمی در بهبود کیفیت و واقع‌نمایی تصاویر خروجی ایفا می‌کنند. این تابع هزینه به صورت ترکیبی از چندین بخش ساخته شده که به مدل امکان می‌دهد تا جزئیات تصویر را حفظ کند و از نظر بصری، تصویر طبیعی و واقعی‌تری ارائه دهد [۲۴, ۳].

در ابتدا، تابع هزینه محتوایی برای حفظ ساختار و اطلاعات اصلی تصویر استفاده می‌شود. این بخش با استفاده از یک شبکه عصبی پیش‌آموزش‌یافته مانند VGG19، ویژگی‌های سطح بالای تصویر اصلی و تصویر بازسازی‌شده را استخراج و مقایسه می‌کند تا از انطباق محتوای اصلی بین این دو اطمینان حاصل شود [۲۵]. در کنار آن، تابع هزینه ادراکی به منظور بهبود کیفیت ادراک شده تصویر افزوده شده است. این قسمت نیز به کمک شبکه VGG19 و لایه‌های میانی آن، ویژگی‌های تصویر بازسازی‌شده و تصویر اصلی را مقایسه می‌کند تا تصویر نهایی از نظر ادراکی تا حد امکان به تصویر واقعی نزدیک باشد و بصری باکیفیت به نظر برسد [۲۶, ۲۷].

همچنین، تابع هزینه تقابل‌گرگشش‌کلیدی در افزایش واقع‌نمایی تصویر بازسازی‌شده دارد. در این بخش، مدل پیشنهادی از یک شبکه تفکیک‌کننده استفاده می‌کند که تلاش می‌کند بین تصاویر واقعی و تصاویر بازسازی‌شده تفاوت قائل شود. در همین حین، شبکه مولد سعی دارد تصاویر بازسازی‌شده‌ای ایجاد کند که بتوانند شبکه تفکیک‌کننده را فریب دهند و از آن عبور کنند، که در نتیجه به بهبود جزئیات و کیفیت بصری تصویر کمک می‌کند [۲۸].

برای محاسبه اختلاف پیکسلی و کاهش خطای کلی، در مدل پیشنهادی به جای توابع مرسوم L1 یا L2، از تابع هزینه Smooth L1 استفاده شده است.

این تابع هزینه به عنوان یک تابع پیوسته و هموار، هم می‌تواند خطاهای کوچک را به صورت جزئی اصلاح کند و هم از تأثیر خطاهای بزرگ به شکل مؤثرتری جلوگیری کند. این ویژگی به مدل کمک می‌کند تا تصاویری با جزئیات دقیق‌تر و کیفیت بالاتر تولید کند و در عین حال نویز کمتری داشته باشد [۲۹].

ترکیب این بخش‌های مختلف در تابع هزینه، موجب می‌شود که مدل پیشنهادی توانایی تولید تصاویری با کیفیت بالا و جزئیات مناسب داشته باشد، در حالی که از لحاظ ادراکی نیز کاملاً طبیعی و نزدیک به تصاویر واقعی به نظر می‌رسند [۲۸].

۳-۱-۳ بلوک توجه کانالی

یکی از مهم‌ترین اجزای شبکه مولد، بلوک توجه کانالی است که با هدف تقویت ویژگی‌های کلیدی در هر کانال طراحی شده است. این بلوک با استفاده از لایه میانگین‌گیر کلی، به طور میانگین ویژگی‌ها را در هر کانال استخراج می‌کند و این اطلاعات را در قالب یک بردار تک‌بعدی کوچک فشرده‌سازی می‌کند. سپس، این بردار از طریق دو لایه کانولوشنی که اندازه آن‌ها به منظور کاهش دهنده ابعاد و افزایش ابعاد تنظیم شده، عبور داده می‌شود.

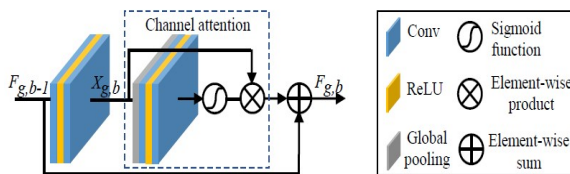
این لایه‌ها با استفاده از یک تابع فعال‌سازی ReLU، ویژگی‌های مهم را تقویت کرده و در نهایت با اعمال یک تابع Sigmoid، وزن‌های جدیدی برای هر کانال محاسبه می‌کنند. خروجی نهایی این بلوک با ورودی اصلی تصویر ترکیب شده و در نتیجه، ویژگی‌های کلیدی تقویت می‌شوند [۵].

۳-۱-۴ بلوک رسوبی توجه کانالی

بلوک رسوبی توجه کانالی یکی از عناصر کلیدی در شبکه عصبی توجه کانال‌محور رسوبی می‌باشد و وظیفه استخراج ویژگی‌های کلیدی و تقویت آن‌ها را برعهده دارد که یکی از عناصر اصلی در بلوک پیشنهادی در مدل پیشنهادی (رسوبی مترام مبتنی بر توجه تقویت شده) است که از ترکیب دو لایه کانولوشنی و بلوک توجه کانالی تشکیل شده است. این بلوک به صورت زیر طراحی شده است:

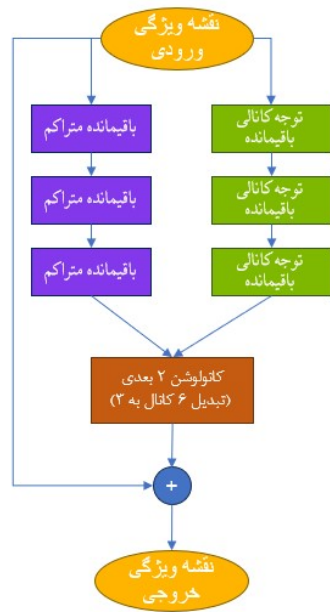
دو لایه کانولوشنی با هسته 3×3 که وظیفه استخراج ویژگی‌های محلی از تصویر را برعهده دارند و تابع فعال‌سازی ReLU که بعد از اولین لایه کانولوشنی قرار گرفته و به تقویت ویژگی‌های استخراج شده کمک می‌کند. بلوک توجه کانالی که پس از استخراج ویژگی‌ها توسط لایه‌های کانولوشنی اعمال می‌شود تا ویژگی‌های کلیدی در هر کانال شناسایی و تقویت شوند.

در نهایت، خروجی این بلوک با ورودی اصلی جمع می‌شود که به حفظ اطلاعات مهم و جلوگیری از ناپدید شدن گرادینان‌ها در طول آموزش کمک می‌کند. ساختار کلی بلوک رسوبی توجه کانالی در شکل ۳ نشان داده شده است [۵].



شکل ۳: ساختار بلوک رسوبی توجه کانالی [۵]

۳-۱-۵ بلوک رسوبی متراکم ۲۵



شکل ۴: ساختار بلوک رسوبی متراکم مبتنی بر توجه تقویت شده

در شکل ۴ ساختار این بلوک نشان داده شده است که ابتدا چندین بلوک رسوبی متراکم قرار دارد که وظیفه استخراج و تقویت ویژگی‌ها را بر عهده دارند. سپس، ویژگی‌های استخراج شده توسط چندین بلوک رسوبی توجه کانالی تقویت و توجه بیشتری به کانال‌های کلیدی داده می‌شود.

در نهایت، خروجی این بلوک‌ها به صورت جمعی ترکیب شده و به یک لایه کانولوشن ۶ به ۳ داده شده تا از میان ویژگی‌های استخراج شده بهترین‌ها را انتخاب یا یک ویژگی با توجه به ویژگی‌های استخراج شده تولید و تعداد کانال‌ها را به تعداد کانال‌های ورودی (۳ کانال) تبدیل و به مرحله بعدی شبکه ارسال کند.

۳-۲ ساختار تفکیک‌کننده

ساختار تفکیک‌کننده مدل پیشنهادی به منظور بهبود تشخیص تصاویر واقعی از تصاویر بازسازی شده طراحی شده است. این تفکیک‌کننده با استفاده از معماری مبتنی بر U-Net و بهره‌گیری از بلوک‌های توجه کانالی، قابلیت شناسایی دقیق‌تر با توجه به جزئیات مهم بین تصاویر واقعی و تولید شده را فراهم کرده است. نوآوری اصلی این بخش در ترکیب لایه‌های کانولوشن با مکانیزم‌های توجه برای تمرکز بر ویژگی‌های محلی و کانالی تصویر و بهبود دقت تفکیک می‌باشد. به‌کارگیری این ساختار باعث افزایش توانایی تفکیک‌کننده در شناسایی جزئیات مهم و بهبود عملکرد کلی شبکه شده است.

بلوک رسوبی متراکم یکی از بلوک‌های اصلی در شبکه عصبی بهبود وضوح تصویر واقعی با هدف بهبود دقت و کارایی در بازسازی تصاویر طراحی شده است. این بلوک شامل پنج لایه کانولوشنی است که به صورت متوالی قرار گرفته‌اند. هر لایه کانولوشنی علاوه بر دریافت خروجی لایه قبلی، ورودی تمامی لایه‌های قبلی را نیز دریافت می‌کند.

این ساختار متراکم به حفظ اطلاعات مهم در طول لایه‌های شبکه کمک کرده و از گم شدن اطلاعات جلوگیری می‌کند.

در نهایت، خروجی نهایی این بلوک با ورودی اولیه جمع می‌شود تا اطلاعات مهم حفظ و جزئیات دقیق‌تر بازسازی شوند. استفاده از این ساختار متراکم، امکان یادگیری ویژگی‌های پیچیده‌تر و عمیق‌تر از تصاویر را فراهم می‌آورد [۲۳].

۳-۱-۶ بلوک رسوبی متراکم مبتنی بر توجه تقویت شده

بلوک رسوبی متراکم مبتنی بر توجه تقویت شده بلوک جدید پیشنهادی است که یکی از اجزای مهم مدل پیشنهادی است، ایده‌آدغام بلوک‌های رسوبی متراکم (در شبکه عصبی بهبود وضوح تصویر واقعی) و بلوک توجه کانالی رسوبی (در شبکه عصبی توجه کانال‌محور رسوبی) از نیاز به بهبود همزمان حفظ جزئیات دقیق و کاهش نویز ناشی شد.

تحلیل نتایج مدل‌های موجود نشان داد که تمرکز صرف بر یک جنبه باعث محدودیت عملکرد کلی می‌شود. در نتیجه، مفهومی برای ترکیب ساختارهای تقویتی پیشنهاد شد که در آن، ارتباطات متراکم برای حفظ و تقویت اطلاعات تصویر و توجه کانالی برای تمرکز بر ویژگی‌های مهم به‌کار گرفته شوند.

برای رسیدن به این ساختار، بررسی‌های مختلف بر روی نحوه جریان اطلاعات و انتقال ویژگی‌ها بین لایه‌ها انجام شد. با تحلیل اثرات انباشتی لایه‌های متراکم و چگونگی بهبود دقت با اضافه کردن توجه در کانال‌های بحرانی، طراحی نهایی شکل گرفت. این فرآیند نتیجه‌ای از آزمایش مکرر پیکربندی‌های مختلف بود که منجر به استفاده از ترکیب دقیق بلوک‌های متراکم و توجه کانالی به‌گونه‌ای شد که قدرت یادگیری بهینه و استحکام بازسازی حفظ گردد. این نوآوری، بازسازی تصاویر با وضوح بالا را با بهبود در معیارهای کیفیت بصری تضمین می‌کند.

Dense Residual Block

Residual Dense Block

Residual Channel Attention Block

۳-۲-۱- بلوک‌های کانولوشنی توجه^{۲۸}

تفکیک‌کننده شامل چندین بلوک کانولوشنی است که در مدل پیشنهادی هر بلوک را توسط بلوک توجه کانالی تقویت شده و وظیفه استخراج ویژگی‌های تصویر را بر عهده دارند. هر بلوک شامل موارد زیر است:

- **لایه کانولوشنی:** این لایه با استفاده از هسته 3×3 و یک تابع فعال‌سازی LeakyReLU ویژگی‌های محلی تصویر را استخراج می‌کند.
- **توجه کانالی:** پس از لایه کانولوشنی، بلوک توجه کانالی اعمال می‌شود تا ویژگی‌های کلیدی در هر کانال تقویت شوند.

۳-۲-۲- تابع هزینه تفکیک‌کننده

تابع هزینه تفکیک‌کننده در مدل پیشنهادی، مشابه شبکه‌ی عصبی بهبود وضوح تصویر و اقیطراحی شده است و به منظور تشخیص تصاویر واقعی از تصاویر تولیدشده توسط مولد، از تابع هزینه باینری متقابل^{۲۹} استفاده می‌کند. این تابع هزینه کمک می‌کند تا تفکیک‌کننده تصاویر واقعی را از تصاویر جعلی تشخیص دهد و همزمان، مولد را مجبور می‌کند تا تصاویر باکیفیت و طبیعی‌تری تولید کند.

این تابع هزینه در رابطه ۱ ارائه شده است:

$$L_D = \sum_{i=1}^N \frac{1}{N} - [y_i \cdot \log(D(x_i)) + (1 - y_i) \cdot \log(1 - D(x_i))] \quad (1)$$

در این فرمول:

- N : تعداد نمونه‌ها در هر دسته است.
- x_i : ورودی i ام به تفکیک‌کننده، که می‌تواند یک تصویر واقعی یا تصویر تولیدشده باشد.
- y_i : برچسب i ام است که برای تصاویر واقعی برابر ۱ و برای تصاویر تولیدی (جعلی) برابر ۰ در نظر گرفته می‌شود.
- $D(x_i)$: خروجی تفکیک‌کننده برای ورودی x_i است که نشان‌دهنده احتمال واقعی بودن تصویر می‌باشد.

هدف این تابع هزینه این است که برای تصاویر واقعی، مقدار $D(x_i)$ را به ۱ و برای تصاویر تولیدشده، آن را به ۰ نزدیک کند. این رویکرد موجب می‌شود تفکیک‌کننده بتواند تفاوت بین تصاویر واقعی و جعلی را تشخیص دهد و همزمان، مولد را به سمت تولید تصاویر باکیفیت‌تر و نزدیک به تصاویر واقعی هدایت کند [۳۰].

۴- مجموعه داده‌ها و معیارهای ارزیابی

در این تحقیق برای ارزیابی و آموزش مدل‌های بهبود کیفیت تصویر، از پایگاه‌های داده معروف DIV2K و OST استفاده شده است. مجموعه داده DIV2K شامل ۱۰۰۰ تصویر با وضوح بالا است که به دو مجموعه آموزش (۸۰۰ تصویر) و اعتبارسنجی/آزمون (۲۰۰ تصویر) تقسیم شده‌اند [۳۱].

این تصاویر شامل مناظر طبیعی، معماری و اشیاء روزمره می‌شوند و به دلیل تنوع بالا و وضوح ۲K، به عنوان منبع استاندارد در مسابقات بین‌المللی مانند NTIRE استفاده می‌شوند. این مجموعه داده به مدل‌های یادگیری عمیق کمک می‌کند تا توانایی تعمیم‌پذیری بهتری داشته باشند [۱۶].

مجموعه داده OST شامل حدود ۳۰۰۰ تصویر با وضوح متنوع است که به طور خاص برای پژوهشگران و توسعه‌دهندگان طراحی شده است. این مجموعه داده شامل تصاویر از موضوعات مختلف از جمله طبیعت، پرتره‌ها، اشیاء شهری و محیط‌های داخلی می‌باشد [۳۲].

تصاویر در این پایگاه به گونه‌ای انتخاب شده‌اند که نمایانگر موقعیت‌ها و شرایط مختلف نوری باشند، و این ویژگی به بهبود الگوریتم‌های یادگیری عمیق و ارزیابی عملکرد آن‌ها کمک می‌کند. با استفاده از این پایگاه‌های داده، توانسته‌ایم مدل‌های بهبود کیفیت تصویر را با دقت و قابلیت تعمیم‌پذیری بیشتری آموزش دهیم [۳۳].

۴-۱- داده افزایشی:

در مدل‌های شبکه عصبی عمیق، داده افزایشی یک مسئله مهم است که می‌تواند تأثیرات قابل توجهی بر کارایی و دقت مدل‌ها داشته باشد. به منظور افزایش تعداد نمونه‌های آموزشی و کاهش حجم داده‌های مورد نیاز برای آموزش، از روش‌های مختلفی برای تولید داده‌های جدید و متنوع استفاده می‌شود. در این بخش، رویکرد مدل پیشنهادی برای مدیریت داده افزایشی با استفاده از مدل تخریب مرتبه بالا و تعریف یک پنجره کانولوشنی شرح داده شده است [۲].

در این پژوهش، از مدل تخریب مرتبه بالا برای ایجاد تنوع در داده‌های آموزشی استفاده شده است. این مدل با اعمال تخریب‌های مختلف به صورت ترکیبی و متوالی، تصاویر جدیدی را ایجاد می‌کند که ویژگی‌های مختلفی دارند.

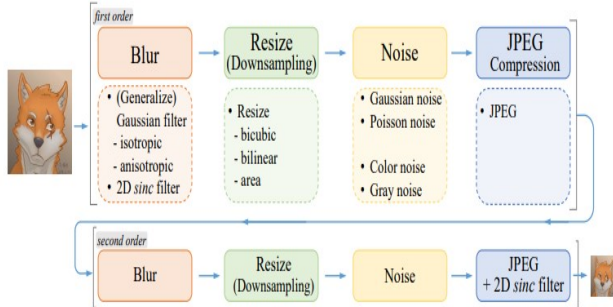
تخریب‌های اعمال شده می‌توانند شامل تاری، کاهش وضوح، نویز و اعوجاج باشند که هر یک به طور جداگانه یا به صورت ترکیبی بر روی تصاویر اعمال می‌شوند و اعمال تخریب‌ها در شکل ۵ نشان داده شده است [۲].

استفاده از پنجره کانولوشنی با هم‌پوشانی مشخص، که نیاز به ذخیره‌سازی و مدیریت حجم زیادی از داده‌ها را کاهش می‌دهد و موجب صرفه‌جویی در فضای ذخیره‌سازی و منابع محاسباتی می‌شود.

رویکرد پیشنهادی برای مدیریت داده افزایشی در مدل‌های شبکه عصبی عمیق، با استفاده از مدل تخریب مرتبه بالا و تعریف پنجره کانولوشنی با هم‌پوشانی مشخص، توانسته است تعداد نمونه‌های آموزشی را افزایش دهد و حجم داده‌های مورد نیاز را کاهش دهد. این روش، با ایجاد تنوع در داده‌های آموزشی و کاهش نیاز به منابع ذخیره‌سازی، به بهبود کارایی و دقت مدل‌های شبکه عصبی کمک شایانی می‌کند [۲].

۴-۲ معیارهای ارزیابی کیفیت تصاویر

برای ارزیابی کیفیت تصاویر بدون نیاز به تصویر مرجع، از معیارهای مختلفی استفاده می‌شود که هر کدام ویژگی‌های خاصی از تصویر را تحلیل می‌کنند. از رایج‌ترین این معیارها می‌توان به BRISQUE، بی‌نظمی، گرادیان، NIQE و PIQE اشاره کرد که به محققان اجازه می‌دهند کیفیت تصاویر را از لحاظ طبیعی بودن، وضوح و نواقص بصری ارزیابی کنند که در جدول ۱ به معرفی این معیارها پرداخته شده است [۷، ۳۷-۳۴].



شکل ۵: مراحل تخریب مرتبه بالا [۲، ۳]

برای تولید داده‌های جدید، یک پنجره کانولوشنی با مقدار مشخصی هم‌پوشانی تعریف شده است. این روش به این صورت عمل می‌کند که تصاویر ورودی با استفاده از پنجره کانولوشنی به بخش‌های کوچکتر تقسیم می‌شوند.

هر پنجره دارای یک هم‌پوشانی مشخص با پنجره‌های مجاور است که موجب ایجاد بخش‌های مشترک بین تصاویر تقسیم‌شده می‌شود. با اعمال پنجره کانولوشنی با هم‌پوشانی مشخص، تعداد بیشتری از بخش‌های تصویر ایجاد می‌شود که به عنوان تصاویر جدید در مجموعه داده‌های آموزشی استفاده می‌شوند.

این رویکرد دارای دو مزیت اصلی است: اول، افزایش تعداد نمونه‌های آموزشی با تقسیم تصاویر ورودی به بخش‌های کوچکتر و اعمال تخریب‌های مختلف، که به بهبود عملکرد مدل و افزایش دقت آن کمک می‌کند؛ و دوم، کاهش حجم داده‌های آموزشی با

جدول ۱: معرفی معیار شناخته شده در زمینه بازسازی کور تصاویر

معیار	توضیح	شاخص بهینگی
NIQE [۳۵, ۷]	ارزیاب کیفیت تصویر طبیعی که بدون مرجع است و به تحلیل ساختار تصویر و مقایسه آن با ویژگی‌های تصاویری که به عنوان طبیعی در نظر گرفته می‌شوند، می‌پردازد.	کمترین مقدار
PIQE [۳۴, ۷]	ارزیاب کیفیت تصویر مبتنی بر ادراک که بدون مرجع است و کیفیت تصویر را در شرایط واقعی و با توجه به ویژگی‌های محلی پیش‌بینی می‌کند.	کمترین مقدار
BRISQUE [۳۶]	ارزیاب کیفیت تصویر بدون مرجع که بر ویژگی‌های آماری تمرکز دارد و کیفیت بصری را بر اساس مدل‌سازی اعوجاج‌های محلی ارزیابی می‌کند.	کمترین مقدار
بی‌نظمی [۳۷]	معیار سنجش میزان اطلاعات یا جزئیات موجود در تصویر که در ارزیابی کیفیت بازسازی تصویر و حفظ جزئیات موثر است؛ هرچه مقدار بیشتر باشد، تصویر اطلاعات بیشتری دارد.	بیشترین مقدار
گرادیان [۳۸]	ارزیاب کیفیت تصویر بر پایه گرادیان که توانایی تصویر در حفظ لبه‌ها و جزئیات ساختاری را می‌سنجد؛ گرادیان بالاتر به معنای لبه‌های تیزتر و جزئیات دقیق‌تر است.	بیشترین مقدار

۵- آزمون و نتایج

در این بخش، به صورت تخصصی به بررسی نتایج به دست آمده از اجرای مدل پیشنهادی و مقایسه آن با مدل‌های پیشین می‌پردازیم. هدف از این بررسی، تحلیل عملکرد مدل در معیارهای PIQE، NIQE، BRISQUE، بی‌نظمی و گرادیان،

این مجموعه از معیارهای بدون مرجع به‌طور کامل جنبه‌های مختلف کیفیت تصاویر بازسازی‌شده را پوشش می‌دهد و به‌خوبی می‌تواند عملکرد مدل پیشنهادی را در مقایسه با سایر مدل‌ها از نظر کیفیت بصری، وضوح، و حفظ جزئیات به دقت ارزیابی کند [۷، ۳۴-۳۸].

جدول ۳: محاسبه متوسط معیار NIQE مدل های پیشین و مقایسه با مدل پیشنهادی

مدل	BSD100	Set5	Set14	Urban100	MANGA109	سال
Bicubic	۵.۷۹	۶.۴۶	۵.۹۴	۵.۴۵	۵.۰۸	۱۹۸۰ به بعد
تقریب بهترین صفحه [۱۱]	۵.۶۲	۶.۲۳	۵.۶۴	۵.۳۲	۴.۸۷	۲۰۱۷
MSRN[۱۷, ۱۸]	۵.۴۳	۶.۱۷	۵.۵۱	۴.۸۱	۴.۲۹	۲۰۱۸
شبکه عصبی عمیق سلسله مراتبی خودتوجه و رمزگذاری بویای تک متغیره [۱۲]	۵.۴۰	۶.۱۶	۵.۴۷	۴.۷۰	۴.۲۰	۲۰۲۳
EDSR[۱۵, ۱۶]	۵.۳۴	۶.۳۰	۵.۴۳	۴.۷۴	۴.۲۳	۲۰۱۷
DRLN[۱۹, ۲۰]	۵.۴۰	۶.۴۱	۵.۵۲	۴.۷۹	۴.۲۹	۲۰۲۰
RCAN[۴, ۵]	۵.۴۰	۶.۴۱	۵.۵۲	۴.۷۹	۴.۲۵	۲۰۱۸
REAL_ESRGAN[۲, ۳]	۳.۵۳	۵.۶۲	۳.۸۱	۳.۵۵	۳.۵۹	۲۰۲۱
A_ESRGAN[۶]	۳.۳۹	۴.۴۹	۳.۲۹	۳.۴۶	۲.۹۶	۲۰۲۲
مدل پیشنهادی	۳.۱۹	۴.۶۱	۳.۴۰	۳.۴۶	۳.۱۴	۲۰۲۵

مقدار نزدیک به کمترین مقدار معیار NIQE را به دست آورده است. برای مجموعه داده BSD100، مقدار ۳/۱۹ ثبت شده که طبیعی تر بودن تصاویر بازسازی شده را نشان می دهد. در مجموعه داده Set5، مقدار ۲/۴۸ به دست آمده که با اختلاف کمی به کمترین مقدار نزدیک بوده و شباهت بالای تصاویر بازسازی شده به نمونه های واقعی را بیان می کند.

در مجموعه داده Set14، مقدار ۳/۴۰ که مشابه دیتاست قبلی نزدیک به کمترین مقدار بوده و عملکرد مطلوب در ایجاد تصاویر طبیعی تر را نشان می دهد. همچنین، در مجموعه داده Urban100 مقدار ۳/۴۶ ثبت شده که توانایی مدل در تولید تصاویر واقع گرایانه تر را تایید می کند.

برای مجموعه داده MANGA109، مقدار ۳/۱۴ ثبت شده که به وضوح نشان دهنده بهبود کیفیت و نزدیکی تصاویر بازسازی شده به تصاویر واقعی می باشد. این مقادیر نشان داده اند که مدل پیشنهادی نه تنها کیفیت بالاتری ارائه می دهد، بلکه تصاویر خروجی آن از نظر بصری به نمونه های طبیعی شباهت بیشتری داشته و احساس واقع گرایانه تری به مخاطب منتقل می کنند [۷, ۳۵].

جدول ۴: محاسبه متوسط معیار BRISQUE مدل های پیشین و مقایسه با مدل پیشنهادی

مدل	BSD100	Set5	Set14	Urban100	MANGA109	سال
Bicubic	۵۳.۸۶	۵۲.۱۵	۵۲.۰۰	۵۱.۳۹	۵۲.۰۰	۱۹۸۰ به بعد
تقریب بهترین صفحه [۱۱]	۴۷.۲۳	۴۸.۰۲	۴۳.۳۱	۴۰.۲۷	۴۵.۰۱	۲۰۱۷
MSRN[۱۷, ۱۸]	۴۳.۹۱	۴۳.۲۰	۴۱.۸۸	۳۹.۹۹	۳۷.۰۵	۲۰۱۸
شبکه عصبی عمیق سلسله مراتبی خودتوجه و رمزگذاری بویای تک متغیره [۱۲]	۴۳.۸۱	۴۲.۰۰	۴۱.۷۴	۳۹.۸۳	۳۶.۷۱	۲۰۲۳
EDSR[۱۵, ۱۶]	۴۳.۶۲	۴۲.۵۴	۴۱.۷۵	۳۹.۷۱	۳۶.۶۷	۲۰۱۷
DRLN[۱۹, ۲۰]	۴۴.۱۱	۴۲.۶۵	۴۱.۸۷	۳۹.۸۲	۳۷.۱۰	۲۰۲۰
RCAN[۴, ۵]	۴۱.۵۵	۴۰.۵۹	۴۰.۸۱	۳۹.۱۸	۳۵.۹۰	۲۰۱۸
REAL_ESRGAN[۲, ۳]	۲۷.۲۲	۳۰.۲۹	۲۷.۶۳	۲۶.۰۲	۲۶.۳۲	۲۰۲۱
A_ESRGAN[۶]	۲۵.۶۷	۲۴.۷۷	۲۳.۸۱	۲۶.۹۰	۲۲.۴۳	۲۰۲۲
مدل پیشنهادی	۱۸.۳۴	۲۷.۴۸	۲۳.۱۱	۲۶.۸۳	۲۳.۴۳	۲۰۲۵

در جدول ۴، معیار BRISQUE مورد ارزیابی قرار گرفته و کیفیت تصاویر بازسازی شده از نظر شباهت به تصاویر واقعی

ارزیابی کیفی تصاویر بازسازی شده، و مقایسه جامع عملکرد با سایر روش های پیشرفته است [۷, ۳۴-۳۸].

این تست ها بر روی چند مجموعه داده معروف در زمینه بهبود کیفیت تصویر انجام شد، از جمله مجموعه داده های BSD100, Urban100, Set14, Set5 و MANGA109 که از استانداردهای شناخته شده برای ارزیابی کیفیت تصاویر محسوب می شوند [۲].

جدول ۲ تا ۴ نتایج کمی معیارهای NIQE، PIQE و BRISQUE را برای مدل های مختلف از جمله مدل پیشنهادی ارائه می دهند که عملکرد برتر این مدل را نسبت به سایر روش ها به وضوح تأیید می کند.

این معیارها بر اساس مقادیری سنجیده می شوند که هرچه کمتر باشند، به معنای کیفیت بهتر تصویر است؛ بنابراین، پایین بودن نمرات در هر سه معیار نشان دهنده عملکرد بهتر مدل در کاهش اعوجاج ها و ارتقای کیفیت بصری است.

جدول ۲: محاسبه متوسط معیار PIQE مدل های پیشین و مقایسه با مدل پیشنهادی

مدل	BSD100	Set5	Set14	Urban100	MANGA109	سال
Bicubic	۸۶.۶۰	۸۷.۳۸	۸۷.۴۴	۸۵.۸۸	۸۶.۲۴	۱۹۸۰ به بعد
تقریب بهترین صفحه [۱۱]	۸۲.۵۶	۸۳.۰۴	۸۲.۳۱	۸۱.۲۲	۸۱.۲۶	۲۰۱۷
MSRN[۱۷, ۱۸]	۷۹.۲۶	۷۹.۹۶	۷۸.۵۸	۷۱.۸۷	۷۳.۸۳	۲۰۱۸
شبکه عصبی عمیق سلسله مراتبی خودتوجه و رمزگذاری بویای تک متغیره [۱۲]	۷۹.۱۳	۷۸.۹۱	۷۷.۸۲	۷۱.۸۰	۷۴.۱۳	۲۰۲۳
EDSR[۱۵, ۱۶]	۷۹.۱۲	۸۲.۱۴	۷۷.۷۱	۷۱.۷۲	۷۴.۱۴	۲۰۱۷
DRLN[۱۹, ۲۰]	۷۹.۴۵	۸۲.۵۰	۷۸.۰۱	۷۲.۰۳	۷۵.۵۱	۲۰۲۰
RCAN[۴, ۵]	۶۱.۶۲	۶۰.۹۲	۶۰.۴۱	۵۴.۲۰	۵۰.۴۷	۲۰۱۸
REAL_ESRGAN[۲, ۳]	۲۸.۰۸	۳۹.۴۱	۳۶.۱۳	۳۷.۲۰	۴۶.۱۵	۲۰۲۱
A_ESRGAN[۶]	۲۵.۰۱	۳۶.۴۶	۳۳.۸۰	۳۲.۱۹	۴۰.۲۵	۲۰۲۲
مدل پیشنهادی	۲۱.۳۹	۲۵.۱۲	۲۰.۷۷	۲۲.۷۳	۳۱.۷۴	۲۰۲۵

معیار PIQE را به دست آورده است. برای مجموعه داده BSD100، مقدار ۲۱/۳۹ ثبت شده است که کاهش مؤثر اعوجاج های بصری و ارتقای کیفیت تصاویر بازسازی شده را نشان می دهد. در مجموعه داده Set5، مقدار ۳۵/۱۲ بیانگر افزایش دقت و بهبود محسوس کیفیت بصری تصاویر می باشد. همچنین، در مجموعه داده Set14 مقدار ۳۰/۷۷ به دست آمده است که کاهش خطاهای ادراکی و تولید تصاویر طبیعی تر را نشان می دهد.

در مجموعه داده Urban100 نیز مقدار ۳۲/۷۳ ثبت شده است که بازسازی واقع گرایانه تر و ارائه تصاویر با جزئیات بیشتر را به نمایش می گذارد. برای مجموعه داده MANGA109، مقدار ۳۱/۷۴ به دست آمده است که اختلاف قابل توجهی نسبت به مقدار ۴۰/۲۵ که بهترین عملکرد مدل های پیشین بوده، داشته است. این مقادیر به وضوح نشان داده اند که مدل پیشنهادی توانسته است در تمامی مجموعه داده ها عملکرد بهتری نسبت به سایر مدل ها ارائه دهد و کیفیت تصاویر بازسازی شده را به طور قابل ملاحظه ای بهبود بخشد [۷, ۳۴].

قابل توجهی بیشتر بوده‌اند. برای مجموعه داده BSD100، مقدار ۷/۳۵ گزارش شده است که نشان می‌دهد تصاویر بازسازی شده اطلاعات بیشتری را حفظ کرده‌اند.

در مجموعه داده Set5، مقدار ۷/۵۱ مشاهده شده است که نشان‌دهنده بالاتر بودن سطح جزئیات در تصاویر بازسازی شده است. مقدار ۷/۵۶ در مجموعه داده Set14 ثبت شده است که بیانگر افزایش کیفیت اطلاعات در تصاویر می‌باشد. همچنین، در مجموعه داده Urban100 مقدار ۷/۶۲ به دست آمده که نشان می‌دهد تصاویر بازسازی شده توانسته‌اند جزئیات بیشتری را بازتاب دهند.

برای مجموعه داده manga109، مقدار ۷/۲۳ ثبت شده که به وضوح اختلاف قابل توجهی با مقدار ۶/۸۸ که بهترین عملکرد مدل‌های قبلی بوده، دارد. این مقادیر به وضوح برتری مدل پیشنهادی را در حفظ اطلاعات و افزایش کیفیت تصاویر بازسازی شده تأیید می‌کنند [۳۴-۳۸, ۷].

جدول ۶: محاسبه متوسط معیار گرادیان مدل‌های پیشین و مقایسه با مدل پیشنهادی

مدل	BSD100	Set5	Set14	Urban100	MANGA109	سال
Bicubic	۲.۶۴	۳.۹۲	۳.۳۴	۳.۸۹	۳.۶۷	۱۹۸۰ به بعد
تقریب بهترین صفحه [۱۱]	۲.۶۴	۴.۰۰	۳.۷۱	۴.۲۱	۴.۳۱	۲۰۱۷
MSRNI [۱۷, ۱۸]	۳.۶۳	۴.۸۰	۴.۳۹	۵.۷۹	۵.۱۹	۲۰۱۸
شبکه عصبی عمیق سلسله مراتبی خودتوجه و رمزگذاری پویای تک متغیره [۱۲]	۳.۶۷	۴.۸۲	۴.۴۲	۵.۸۹	۵.۲۶	۲۰۲۳
EDSR [۱۵, ۱۶]	۳.۶۹	۴.۸۵	۴.۴۷	۶.۰۰	۵.۳۰	۲۰۱۷
DRLNI [۱۹, ۲۰]	۳.۶۶	۴.۸۲	۴.۴۵	۵.۷۸	۵.۲۷	۲۰۲۰
RCAN [۴, ۵]	۳.۶۵	۴.۹۳	۴.۴۳	۵.۵۰	۵.۱۷	۲۰۱۸
REAL_ESRGAN [۲, ۳]	۴.۵۶	۵.۸۵	۵.۲۵	۶.۹۰	۵.۱۶	۲۰۲۱
A_ESRGAN [۶]	۴.۸۱	۶.۳۲	۵.۶۳	۷.۱۳	۵.۴۳	۲۰۲۲
مدل پیشنهادی	۵.۹۱	۶.۱۳	۶.۳۱	۷.۵۱	۶.۲۱	۲۰۲۵

در جدول ۶، معیار گرادیان مورد بررسی قرار گرفته است که به سنجش تغییرات شدت روشنایی و جزئیات لبه‌های تصاویر بازسازی شده می‌پردازد. مقادیر ثبت شده برای مدل پیشنهادی در مقایسه با سایر مدل‌ها به طور قابل توجهی بیشتری با اختلاف کم نزدیک بوده‌اند. برای مجموعه داده BSD100، مقدار ۵/۹۱ گزارش شده است که نشان می‌دهد تصاویر بازسازی شده کیفیت بالاتری از جزئیات را حفظ کرده‌اند. در مجموعه داده Set5، مقدار ۶/۱۲ مشاهده شده است که حاکی از افزایش چشمگیر در حفظ جزئیات لبه‌ها در تصاویر بازسازی شده است. مقدار ۶/۳۱ در مجموعه داده Set14 ثبت شده است که نمایانگر افزایش کیفیت در بازسازی تصاویر است. همچنین، در مجموعه داده Urban100 مقدار ۷/۵۱ به دست آمده که بیانگر توانایی مدل در بازسازی جزئیات پیچیده و دقیق تصاویر می‌باشد. برای مجموعه داده manga109، مقدار ۶/۲۱ به ثبت رسیده است که به وضوح اختلاف قابل توجهی با مقدار ۵/۴۲ که بهترین عملکرد مدل‌های

بررسی شده است. مقادیر به دست آمده توسط مدل پیشنهادی به طور قابل توجهی کمتر یا نزدیک به کمترین مقدار از سایر مدل‌ها گزارش شده‌اند. برای مجموعه داده BSD100، مقدار ۱۸/۳۴ ثبت شده که بهبود کیفیت تصاویر بازسازی شده را نشان می‌دهد. در مجموعه داده Set5، مقدار ۲۷/۴۸ گزارش شده و شباهت بیشتر تصاویر بازسازی شده به نمونه‌های واقعی را تأیید می‌کند. مقدار ۲۳/۱۱ در مجموعه داده Set14 به ثبت رسیده که کیفیت بالاتر تصاویر بازسازی شده را تأیید می‌کند.

همچنین، مقدار ۲۶/۸۳ در مجموعه داده Urban100 گزارش شده که بازسازی تصاویر با جزئیات بیشتر و شباهت بیشتر به تصاویر واقعی را نشان می‌دهد. برای مجموعه داده manga109، مقدار ۲۳/۴۳ ثبت شده که اختلاف نزدیکی با مقدار ۲۲/۴۳ بهترین عملکرد مدل‌های قبلی دارد و برتری مدل پیشنهادی در تولید تصاویر طبیعی و با کیفیت بالا را تأیید می‌کند [۳۶].

نتایج جداول ۲ تا ۴ به خوبی نشان می‌دهند که مدل پیشنهادی در هر سه معیار عملکردی برتر و یا در برخی موارد نزدیک به بهترین مدل‌ها داشته است و می‌تواند در مقایسه با روش‌های دیگر، تصاویری با کیفیت بالاتر و وضوح بهتر ارائه دهد. این یافته‌ها تأکید می‌کنند که مدل پیشنهادی گزینه‌ای مؤثر و قابل اعتماد در حوزه بازسازی تصاویر با کیفیت بصری بالا است و می‌تواند نیازهای متنوع در پردازش تصویر را به خوبی برآورده سازد [۳۶-۳۴].

جداول ۵ و ۶ در کنار جداول قبلی (۲ تا ۴) تکمیل‌کننده ارزیابی جامع از عملکرد مدل پیشنهادی هستند و کیفیت بالای این مدل را در مقایسه با سایر مدل‌ها به خوبی نشان می‌دهند. این جداول، نتایج کمی معیارهای بی نظمی و گرادیان را برای مدل‌های مختلف از جمله مدل پیشنهادی ارائه می‌دهند که در این دو معیار، مقادیر بالاتر نشان‌دهنده کیفیت بهتر تصویر است.

جدول ۵: محاسبه متوسط معیار بی نظمی مدل‌های پیشین و مقایسه با مدل پیشنهادی

مدل	BSD100	Set5	Set14	Urban100	MANGA109	سال
Bicubic	۷.۲۳	۷.۵۷	۷.۴۴	۷.۵۷	۷.۰۱	۱۹۸۰ به بعد
تقریب بهترین صفحه [۱۱]	۷.۲۳	۷.۵۲	۷.۴۳	۷.۵۷	۶.۹۹	۲۰۱۷
MSRNI [۱۷, ۱۸]	۷.۲۴	۷.۵۱	۷.۴۳	۷.۵۶	۶.۹۲	۲۰۱۸
شبکه عصبی عمیق سلسله مراتبی خودتوجه و رمزگذاری پویای تک متغیره [۱۲]	۷.۲۳	۷.۴۸	۷.۴۳	۷.۵۶	۶.۹۲	۲۰۲۳
EDSR [۱۵, ۱۶]	۷.۲۴	۷.۵۰	۷.۴۴	۷.۵۶	۶.۹۳	۲۰۱۷
DRLNI [۱۹, ۲۰]	۷.۲۴	۷.۴۹	۷.۴۳	۷.۵۶	۶.۹۱	۲۰۲۰
RCAN [۴, ۵]	۷.۲۶	۷.۵۹	۷.۴۹	۷.۶۰	۷.۰۹	۲۰۱۸
REAL_ESRGAN [۲, ۳]	۷.۲۱	۷.۴۹	۷.۴۳	۷.۵۳	۶.۶۹	۲۰۲۱
A_ESRGAN [۶]	۷.۲۵	۷.۴۵	۷.۴۹	۷.۵۷	۶.۸۸	۲۰۲۲
مدل پیشنهادی	۷.۲۵	۷.۵۱	۷.۵۶	۷.۶۲	۷.۲۳	۲۰۲۵

اطلاعات موجود در تصاویر بازسازی شده می‌پردازد. مقادیر ثبت شده برای مدل پیشنهادی در مقایسه با سایر مدل‌ها به طور

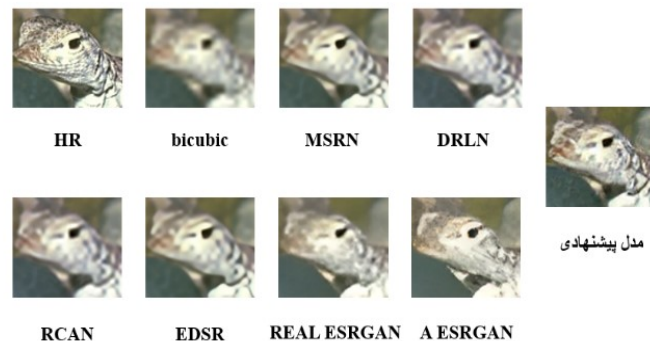
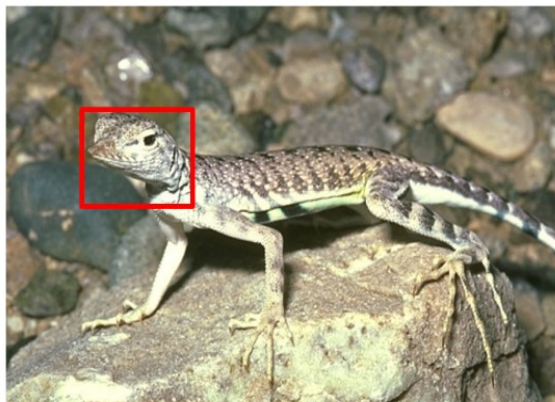
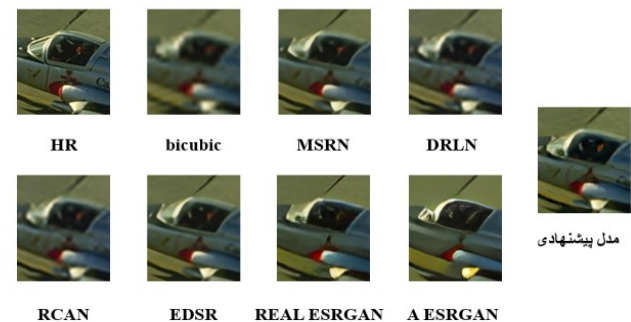
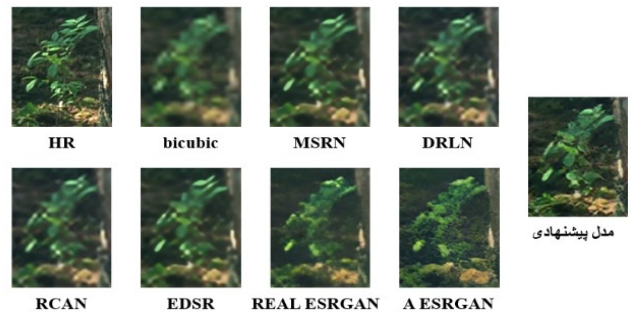
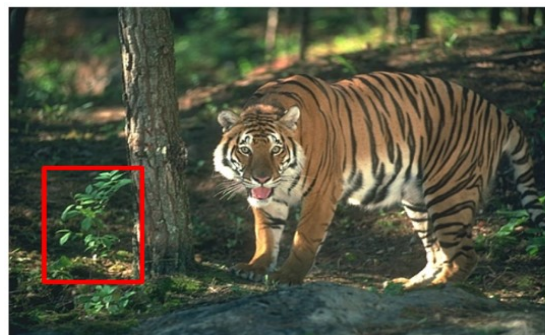
و لبه‌های تصویر به‌وضوح بیشتر و دقیق‌تر دیده می‌شوند، در حالی که مدل‌های دیگر اغلب با مشکلاتی مانند تاری یا عدم دقت در بازسازی جزئیات مواجه هستند.

مقایسه این تصاویر در کنار نتایج کمی جداول ۲ تا ۶، برتری مدل پیشنهادی را در حفظ ساختارهای تصویری و ارائه کیفیتی برتر به‌خوبی نشان می‌دهد. ترکیب این نتایج و تصاویر به‌وضوح تأیید می‌کند که مدل پیشنهادی در بازسازی تصاویر با وضوح بالا و کیفیت بصری بهتر، موفق عمل کرده و می‌تواند گزینه‌ای مطمئن و قدرتمند در پردازش و بازسازی تصویر در کاربردهای مختلف باشد.

قبلی بوده، دارد. این مقادیر به روشنی برتری مدل پیشنهادی را در بازسازی تصاویر با حفظ کیفیت و جزئیات لبه‌ها تأیید می‌کنند [۷، ۳۴-۳۸].

مجموعه شکل ۵ شامل تصاویر بازسازی‌شده توسط مدل‌های مختلف از جمله مدل پیشنهادی را نمایش می‌دهند که تفاوت‌های کیفی میان خروجی مدل پیشنهادی و سایر مدل‌ها را به‌خوبی به تصویر می‌کشند.

این تصاویر نشان می‌دهند که مدل پیشنهادی، علاوه بر ارائه کیفیت بصری بالاتر، جزئیات بیشتری را نیز حفظ کرده و توانسته است وضوح و دقت بهتری را نسبت به مدل‌های دیگر فراهم آورد. در تصاویر بازسازی‌شده توسط مدل پیشنهادی، جزئیات ریز



شکل ۵: مقایسه کیفیت خروجی مدل‌های مختلف بر مجموعه تصاویر

بازسازی‌شده به کار گرفته شده‌اند. عملکرد مدل پیشنهادی در تمامی این معیارها بهبود قابل توجهی نسبت به مدل‌های پیشین نشان داده است.

۶- نتیجه‌گیری و پیشنهادات برای کارهای آینده

در این بخش، نتایج مدل پیشنهادی تحلیل و ارزیابی شده است. معیارهای مختلف از جمله NIQE، PIQE، BRISQUE، بی‌نظمی و گرادیان برای ارزیابی کیفیت تصاویر

این مدل نه تنها در حذف نویز، بلکه در بازسازی جزئیات دقیق نیز موفق بوده و کارایی خود را در برابر مدل‌های پیشین ثابت کرده است. کاربردهای بالقوه این مدل شامل تصاویر پزشکی، ماهواره‌ای، نظارتی و بهبود تصاویر قدیمی و ... است.

۶-۲ پیشنهادات برای کارهای آینده.

- بهینه‌سازی بیشتر پارامترهای شبکه: بررسی و بهینه‌سازی پارامترهایی مانند تعداد لایه‌ها و مقادیر هاپیرپارامترها می‌تواند به عملکرد بهتر مدل در مواجهه با تصاویر متنوع کمک کند.
- استفاده از داده‌های آموزشی گسترده‌تر و متنوع‌تر: افزودن مجموعه داده‌های جدید با انواع شرایط نوری، نویز و اعوجاج‌های مختلف، می‌تواند به مدل کمک کند تا توانایی تعمیم‌پذیری بالاتری داشته باشد و در کاربردهای واقعی عملکرد مطلوبی ارائه دهد.
- ترکیب مدل با روش‌های نوین یادگیری عمیق: ترکیب این مدل با تکنیک‌های جدیدی مانند شبکه‌های مولد متخاصم چندگانه (یا شبکه‌های بازگشتی پیچشی^{۳۳}) می‌تواند به بهبود بازسازی جزئیات و افزایش دقت در شرایط پیچیده کمک کند.
- ارزیابی عملکرد مدل در کاربردهای خاص: آزمایش مدل در حوزه‌های خاص مانند تصاویر پزشکی و ماهواره‌ای می‌تواند به توسعه آن برای شرایط کاربردهای ویژه کمک کند. این امر موجب می‌شود که مدل پیشنهادی به ابزاری تخصصی‌تر و مؤثرتر برای تحلیل و بازسازی تصاویر در این حوزه‌ها تبدیل شود.

با توجه به توانایی‌های مدل پیشنهادی در بهبود کیفیت تصاویر و بازسازی جزئیات، این تحقیقات آینده می‌تواند به توسعه بیشتر و استفاده مؤثرتر از این مدل در کاربردهای مختلف بینایی کامپیوتری کمک کنند.

۷- منابع:

- [۱] M. Impieri, "Methods for Blind Super-Resolution of satellite images," Politecnico di Torino, 2023.
- [۲] Z. Zhu, Y. Lei, Y. Qin, C. Zhu, and Y. Zhu, "IRE: improved image super-resolution based on real-ESRGAN," IEEE Access, vol. 11, pp. 45334-45348, 2023.

در معیار PIQE، مدل پیشنهادی با مقدار ۲۱/۳۹ برای مجموعه داده BSD100 و مقدار ۳۱/۷۴ برای مجموعه داده manga109 عملکردی بهینه داشته است. این مقادیر نشان‌دهنده کاهش مؤثر اعوجاج‌های بصری و بهبود کیفیت تصاویر بازسازی‌شده می‌باشند. در معیار NIQE، مقدار ۳/۱۹ در مجموعه داده BSD100 و ۳/۸۴ در مجموعه داده manga109 به دست آمده است که نشان‌دهنده شباهت بیشتر تصاویر بازسازی‌شده به تصاویر طبیعی است. این نتایج حاکی از توانایی مدل در تولید تصاویر طبیعی‌تر و واقع‌گرایانه‌تر هستند.

در معیار BRISQUE، مدل پیشنهادی با مقدار ۱۸/۳۴ برای مجموعه داده BSD100 و مقدار ۲۷/۴۸ برای مجموعه داده manga109 کیفیت بصری برتری را بدون نیاز به تصویر مرجع به نمایش گذاشته است. این مقادیر بیانگر افزایش شباهت به نمونه‌های واقعی می‌باشند. در معیار بی‌نظمی، مقدار ۷/۳۵ برای مجموعه داده BSD100 و مقدار ۷/۲۳ برای مجموعه داده manga109 ثابت شده است که نشان می‌دهد تصاویر بازسازی‌شده اطلاعات و جزئیات بیشتری را حفظ کرده‌اند.

در معیار گرادیان، مقدار ۵/۹۱ برای مجموعه داده BSD100 و مقدار ۶/۲۱ برای مجموعه داده manga109 مشاهده شده است که نشان‌دهنده بهبود قابل ملاحظه در حفظ لبه‌ها و ساختارهای دقیق تصاویر بازسازی‌شده است. این نتایج به وضوح تأیید می‌کنند که استفاده از بلوک‌های ترکیبی و بهینه‌سازی‌های معرفی‌شده به بهبود همزمان دقت جزئیات و کاهش نویز منجر شده است. مقایسه‌ها در جداول نیز برتری مدل پیشنهادی را در برابر شبکه بهبود وضوح تصویر با تقویت توجه شبکه بهبود وضوح تصویر واقعی مشخص می‌کنند.

۶-۱ نتیجه‌گیری و جمع‌بندی

یک روش نوآورانه برای بازسازی تصاویر با وضوح بالا و حذف نویزهای پیچیده در این مقاله معرفی شده است. با ترکیب قابلیت‌های شبکه توجه کانال محور رسوبی شبکه بهبود وضوح تصویر واقعی، نواقص هر کدام برطرف شده و بهبود قابل توجهی در کیفیت تصاویر بازسازی‌شده حاصل گردیده است. این ترکیب به واسطه بهره‌گیری از توجه کانالی برای حفظ جزئیات و استفاده از تخریب‌های تصادفی برای طبیعی‌تر کردن تصاویر امکان‌پذیر شد. تغییرات کلیدی شامل معرفی بلوک رسوبی متراکم مبتنی بر توجه تقویت شده، و بهینه‌سازی توابع هزینه بوده است. نتایج نشان داد که مدل پیشنهادی در شرایط بازسازی کور و معیارهای PIQE، NIQE و ... برتری چشم‌گیری دارد. به طور کلی، ترکیب بهینه روش‌های توجه کانالی و کاهش نویز به بهبود کیفیت بصری و حفظ ساختارهای جزئی کمک کرده است.

- [۱۶] C. Fiscone et al., "Generalizing the Enhanced-Deep-Super-Resolution neural network to brain MR images: a retrospective study on the Cam-CAN dataset," *Eneuro*, vol. 11, no. 5, 2024.
- [۱۷] H. Wang, "MSRDN: A Super-Resolution Network for Human Body," in *2024 3rd International Conference on Innovations and Development of Information Technologies and Robotics (IDITR)*, 2024: IEEE, pp. ۱۷۸-۱۸۲.
- [۱۸] D. Kong, L. Gu, X. Li, and F. Gao, "Multi-Scale Residual Dense Network for the Super-Resolution of Remote Sensing Images," *IEEE Transactions on Geoscience and Remote Sensing*, 2024.
- [۱۹] S. Anwar and N. Barnes, "Densely residual laplacian super-resolution," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 3, pp. 1192-1۲۰۴, ۲۰۲۰.
- [۲۰] I. Sayal, "Image transformation using super resolution," ۲۰۲۴.
- [۲۱] I. J. Goodfellow, "On distinguishability criteria for estimating generative models," *arXiv preprint arXiv:1412.6515*, 2014.
- [۲۲] I. Goodfellow et al., "Generative adversarial networks," *Communications of the ACM*, vol. 63, no. 11, pp. 139-۱۴۴, ۲۰۲۰.
- [۲۳] X. Wang et al., "Esrgan: Enhanced super-resolution generative adversarial networks," in *Proceedings of the European conference on computer vision (ECCV) workshops, 2018*, pp. 0-0.
- [۲۴] X. Li and S. He, "Blind image quality evaluation method based on cyclic generative adversarial network," *IEEE Access*, 2024.
- [۲۵] L. Zhang, L. Wang, Z. Yan, Z. Jia, H. Wang, and X. Tang, "Star Generative Adversarial VGG Network-Based Sample Augmentation for Insulator Defect Detection," *International Journal of Computational Intelligence Systems*, vol. 17, no. 1, pp. 1-1۶, ۲۰۲۴.
- [۲۶] J. Lin et al., "Improving image restoration through removing degradations in textual representations," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 2866-2۸۷۸.
- [۲۷] Z. Zhang et al., "Rethink arbitrary style transfer with transformer and contrastive learning," *Computer Vision and Image Understanding*, vol. 241, p. 103951, 2024.
- [۲۸] E. Gocer, "GANbased augmentation using a hybrid loss function for dermoscopy images," *Artificial Intelligence Review*, vol. 57, no. 9, p. 234, 2024.
- [۲۹] Y. Li, F. Xiao, W. Liang, and L. Gui, "Multiply complementary priors for image compressive sensing reconstruction in impulsive noise," *ACM Transactions on Multimedia Computing, Communications and Applications*, vol. 20, no. 6, pp. 1-2۲, ۲۰۲۴.
- [۳۰] G. George, "Pretrained U-Net: in-depth analysis of binary image segmentation in underwater marine
- [۳] R. Li and W. Zhou, "Image Super-Resolution Reconstruction of Landslide Based on Real-ESRGAN," in *Proceedings of the 2024 8th International Conference on Control Engineering and Artificial Intelligence*, 2024, pp. 203-208.
- [۴] Z. Huang, "Research on Super-Resolution Processing of Improved RCAN on SMT Blurred Images," in *2024 7th International Conference on Artificial Intelligence and Big Data (ICAIBD)*, 2024: IEEE, pp. 478-4۸۳.
- [۵] Z. Lin et al., "Revisiting rcnn: Improved training for image super-resolution," *arXiv preprint arXiv:2201.11279*, 2022.
- [۶] Z. Wei, Y. Huang, Y. Chen, C. Zheng, and J. Gao, "A-ESRGAN: Training real-world blind super-resolution with attention U-Net Discriminators," in *Pacific Rim International Conference on Artificial Intelligence*, 2023: Springer, pp. 16-27.
- [۷] S. Higashiyama et al., "Usefulness of a No-Reference Metric for Evaluation of Images in Nuclear Medicine-A Comparative Study with Visual Assessment," 2021.
- [۸] J. Lin et al., "Suppressing Uncertainties in Degradation Estimation for Blind Super-Resolution," *arXiv preprint arXiv:2406.16459*, 2024.
- [۹] A. H. Khan, C. Micheloni, and N. Martinel, "IDENet: Implicit degradation estimation network for efficient blind super resolution," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 6065-6۰۷۵.
- [۱۰] X. Wang, L. Sun, A. Chehri, and Y. Song, "A review of GAN-based super-resolution reconstruction for optical remote sensing images," *Remote Sensing*, vol. 15, no. 20, p. 5062, 2023.
- [۱۱] ع. م. اسمعیلی زینی، ع. م. لطیف، و ق. برید لقمانی، "بزرگ‌نمایی تصویر" *vol. 5, no. 1, pp. ۲۶-۳۵, ۲۰۲۱.*
- [۱۲] ع. دلشادی، و. مهرداد، و م. دولت‌شاهی، "بهبودسازی ساختاری شبکه عصبی عمیق سلسله مراتبی خودتوجه و رمزگذاری پویای تک متغیره برای بزرگ‌نمایی تصاویر دیجیتال" *مجله بینایی ماشین و پردازش تصویر*, *vol. ۱۰, no. ۳, pp. ۴۷-۶۳, ۲۰۲۳.*
- [۱۳] C. M. Ward, J. Harguess, B. Crabb, and S. Parameswaran, "Image quality assessment for determining efficacy and limitations of Super-Resolution Convolutional Neural Network (SRCNN)," in *Applications of Digital Image Processing XL*, 2017, vol. ۱۰۳۹۶: SPIE, pp. ۱۹-۳۰.
- [۱۴] D. Vint, G. Di Caterina, J. Soraghan, R. Lamb, and D. Humphreys, "Evaluation of performance of VDSR super resolution on real and synthetic images," in *2019 Sensor Signal Processing for Defence Conference (SSPD)*, 2019: IEEE, pp. 1-5.
- [۱۵] B. M. Kuriakose, "EDSR: Empowering super-resolution algorithms with high-quality DIV2K images," *Intelligent Decision Technologies*, vol. 17, no. 4, pp. 1249-1۲۶۳, ۲۰۲۳.

- environment," *Machine Learning for Computational Science and Engineering*, vol. 1, no. 1, p. 4, 2025.
- [۳۱] K. Huang, X. Yin, T. Gu, and W. Gao, "Perceptual-Centric Image Super-Resolution using Heterogeneous Processors on Mobile Devices," 2024.
- [۳۲] S. Wu and J. Zhang, "TargetSR: Towards Semantic location Real-World Image Super-Resolution with Diffusion Prior," in *2024 International Joint Conference on Neural Networks (IJCNN)*, 2024: IEEE, pp. 1-8.
- [33] <https://www.kaggle.com/datasets/thaihoa1476050/df2k-ost>.
- [۳۴] J. Greffier et al., "Comparison of two deep-learning image reconstruction algorithms on cardiac CT images: A phantom study," *Diagnostic and Interventional Imaging*, vol. 105, no. 3, pp. 110-117, 2024.
- [۳۵] G. Omarova et al., "No-reference quality assessment of medical images using contrast enhancement," *Journal of Theoretical and Applied Information Technology*, vol. 101, no. 1, pp. 267-281, 20۲۳.
- [۳۶] X. Li, J. Pan, J. Shang, A. Souri, and M. Gao, "An improved blind/referenceless image spatial quality evaluator algorithm for image quality assessment," *International Journal of Computational Science and Engineering*, vol. 27, no. 1, pp. 48-5۶, ۲۰۲۴.
- [۳۷] A. Gavrovska, A. Samčović, and D. Dujković, "No-Reference Image Quality Assessment Based on Machine Learning and Outlier Entropy Samples," *Pattern Recognition and Image Analysis*, vol. 34, no. 2, pp. 275-۲۸۷, ۲۰۲۴.
- [۳۸] H. F. Tolie and M. R. Faraji, "Screen content image quality assessment using distortion-based directional edge and gradient similarity maps," *Signal Processing: Image Communication*, vol. 101, p. 116562, 2022.