

مروری بر شبکه‌های عصبی حذف نویز تصویر

جواد آسلیمی ضامنجان^۱، محمدحسین شکور^۲، محسن رحمانی^۳

چکیده

حذف نویز تصویر یک موضوع فعال در پردازش تصویر سطح پایین است و تاکنون تحقیقات بسیاری در این زمینه انجام شده است. روش‌های حذف نویز تصویر را می‌توان به دو دسته کلی روش‌های مبتنی بر مدل و روش‌های مبتنی بر یادگیری افتراقی تقسیم‌بندی کرد. شبکه‌های عصبی کانولوشن (CNN) عمیق یکی از روش‌های مبتنی بر یادگیری هستند که در سال‌های اخیر نتایج خوبی را در وظایف مختلف بازیابی تصویر (از جمله حذف نویز تصویر) ارائه کردند. در این پژوهش ابتدا روش‌های حذف نویز مبتنی بر مدل به صورت اجمالی معرفی و مزایا و معایب آن‌ها بررسی شده است؛ سپس شبکه‌های عصبی حذف نویز، از نظر معماری شبکه و بخش‌های تشکیل دهنده، مورد بررسی قرار گرفته و ساختارهای مورد استفاده در این شبکه‌ها، که موجب بهبود عملکرد حذف نویز شده است، معرفی شده‌اند. در بخش پایانی این مقاله ساختار کلی مشترک شبکه‌های عصبی حذف نویز مدرن، به همراه عملکرد بخش‌های مختلف این ساختار مورد بررسی قرار گرفته است. این مقاله می‌تواند مرجع مناسبی برای مطالعه ایده‌های مورد استفاده در معماری شبکه‌های عصبی کانولوشن حذف نویز باشد.

کلید واژه‌ها

حذف نویز تصویر، شبکه عصبی عمیق، نویز گوسی، توصیفگر تصویر، بازیابی تصویر

ایجاد نتایج نامناسبی مانند: اعوجاج (یا تغییر) پیکسل‌ها، ایجاد خطوط ناهموار، ایجاد لبه‌های غیر واقعی، حذف خطوط و گوشه‌ها، ایجاد مصنوعات و اشیاء مات در تصویر شود. این سیگنال ناخواسته می‌تواند به دلایل مختلف از جمله پدیده‌های طبیعی، استفاده از تجهیزات معیوب، فشرده سازی و خطاهای انتقال ایجاد شود [۱].

انواع مختلفی از نویز می‌توانند بر روی یک تصویر تأثیر بگذارند و موجب از بین رفتن اطلاعات مفید آن شوند. برخی از این نویزها عبارتند از: نویز گوسی^۱، نویز نمک و فلفل^۲، نویز لکه‌ای^۳، نویز

۱- مقدمه

نویز یک سیگنال ناخواسته است که در مقادیر رنگ یا شدت روشنایی یک تصویر تغییرات تصادفی ایجاد می‌کند و موجب از بین رفتن اطلاعات ارزشمند تصویر می‌شود. نویز می‌تواند باعث

این مقاله در خردادماه ۱۴۰۳ دریافت و در مردادماه پذیرفته گردید.

^۱ دانشجوی دکتری مهندسی کامپیوتر-نرم افزار، دانشکده فنی و مهندسی، گروه مهندسی کامپیوتر، دانشگاه اراک، اراک، ایران.

رایانامه: javad.asalimi@gmail.com

^۲ استادیار، دانشکده فنی و مهندسی، گروه مهندسی کامپیوتر، دانشگاه اراک، اراک، ایران.

رایانامه: mh-shakoor@araku.ac.ir

^۳ دانشیار، دانشکده فنی و مهندسی، گروه مهندسی کامپیوتر، دانشگاه اراک، اراک، ایران.

رایانامه: m-rahmani@araku.ac.ir

نویسنده مسئول: محمدحسین شکور

^۱Unrealistic Edges

^۲Artifacts

^۳Gaussian Noise

^۴Salt and Pepper Noise

می‌شوند و اطلاعات ارزشمندی را در زمینه‌های مختلف ارائه می‌دهند. اغلب این تصاویر در شرایط جوی نامناسب و یا نور ضعیف ثبت شده‌اند از این رو روش‌های بازیابی و حذف نویز تصویر به ابزار ضروری برای تجزیه و تحلیل و پردازش تصویر در تمام این زمینه‌ها تبدیل شده‌اند [۴].

۱-۱- کاربردهای عملی حذف نویز تصویر

حذف نویز تصویر در بخش‌های مختلف زندگی، از جمله نظامی، صنعتی، پزشکی، سنجش از دور و رباتیک کاربرد دارد. در این بخش به بررسی اجمالی کاربرد حذف نویز تصویر در انواع مختلف تصویربرداری پزشکی^{۱۱}، تصاویر ماهواره‌ای، رادارهای روزنه مصنوعی^{۱۲} و دوربین‌های نظارت تصویری^{۱۳} می‌پردازیم.

روش‌های مختلف تصویربرداری پزشکی در دهه‌های اخیر پیشرفت‌های سریعی داشته‌اند و به بخش بسیار مهمی از فرآیند تشخیص بیماری تبدیل شده‌اند. تصاویر پزشکی نقش حیاتی در به دست آوردن اطلاعات از قلب، مغز، اعصاب و سایر اعضای داخلی بدن انسان دارند. با اعمال پردازش‌های مختلف بر روی تصاویر پزشکی می‌توان بافت سالم و آسیب دیده بدن را از یکدیگر تفکیک کرد؛ با این حال، از دست دادن اطلاعات هر ناحیه از بدن در طول فرآیند تصویربرداری پزشکی ممکن است منجر به فاجعه‌ای مانند مرگ شود. چالش اصلی در فرآیند تصویربرداری پزشکی، دستیابی به تصویر اعضای بدن، بدون از دست دادن کوچک‌ترین بخش از اطلاعات مهم آن است. یکی از مهم‌ترین مشکلات مهم در بکارگیری تصاویر پزشکی، تخریب تصویر به وسیله نویز یا سایر مصنوعات در طول فرآیند دریافت تصویر یا مراحل پردازشی بعدی است. زمانی اهمیت نویز افزایش پیدا می‌کند که اشیاء تصویربرداری شده کوچک و دارای کنتراست نسبتاً کمی باشند. در مقایسه با تصاویر معمولی، تصاویر پزشکی کنتراست کمی دارند. بنابراین حذف نویز از تصاویر پزشکی بسیار حیاتی است زیرا این نویزها ممکن است شناسایی بیماری‌ها را مختل کنند [۷].

روش‌های رایج تصویربرداری پزشکی عبارتند از تصاویر مافوق صوت^{۱۴} (US) یا سونوگرافی، تصاویر تشدید مغناطیسی^{۱۵} (MR)، تصاویر توموگرافی کامپوتی^{۱۶} یا سی تی اسکن (CT) و تصاویر توموگرافی انتشار پوزیترون^{۱۷} (PET). تصویربرداری سونوگرافی (US) یکی از پرکاربردترین روش‌های تصویربرداری پزشکی در زمینه قلب و عروق و زنان و زایمان است. محبوبیت سونوگرافی به این دلیل است که تصاویر با وضوح بالا را بدون استفاده از تابش یونیزه ارائه می‌دهد. در تصویربرداری سونوگرافی

ضربه‌ای^۱، نویز پواسون^۲ [۲]. رایج‌ترین نویزی که موجب تخریب تصاویر می‌شود، نویز گوسی سفید جمع‌شونده^۳ (AWGN) است. این نویز به طور گسترده برای مدل‌سازی نویز حرارتی استفاده می‌شود [۳]. این نویز از نوع نویزهای آماری است که دارای تابع چگالی احتمال^۴ (PDF) مشابه با توزیع گوسی (توزیع نرمال) است. منبع اصلی نویز گوسی در زمان تصویر برداری است. این نویز عمدتاً در هنگام دریافت تصویر و در مدارهای آنالوگ تجهیزات تصویر برداری ایجاد می‌شود. این نویز به عنوان قسمت اصلی نویز در یک سنسور تصویربرداری شناخته می‌شود، به همین دلیل به آن نویز سنسور نیز گفته می‌شود. نویز سنسور تصویر برداری در تصویربرداری با نور کم یا دمای بالا نویز گوسی است [۲، ۴].

حذف نویز تصویر یک موضوع کلاسیک و در عین حال فعال در پردازش تصویر سطح پایین^۵ است و تاکنون تحقیقات بسیاری در این زمینه انجام شده است. هدف از حذف نویز تصویر، بازیابی تصویر تمیز^۶ یا سالم (x) از نمایش نویزی تصویر (y) است. تصویر نویزی از مدل تخریب رابطه (۱) پیروی می‌کند. یک فرض بسیار رایج این است که n نویز گوسی سفید جمع‌شونده (AWGN) با انحراف معیار استاندارد σ است [۵].

$$y = x + n(1)$$

هدف از حذف نویز تصویر^۷، بهبود تفسیرپذیری یک تصویر با حذف اعوجاج‌های نامطلوب یا بهبود ویژگی‌های تصویر، برای پردازش‌های بعدی است. حذف نویز تصویر برای بیشتر کاربردهای پردازش تصویر، یک مرحله پیش پردازش است. حذف نویز تصویر در وظایف مختلف پردازش تصویر مانند: بازیابی تصویر، ردیابی بصری^۸، بخش‌بندی تصویر^۹، طبقه‌بندی تصویر^{۱۰} و ... کاربرد دارد؛ از این رو هرگونه پیشرفت در حوزه حذف نویز تصویر موجب بهبود عملکرد وظایف مختلف تجزیه و تحلیل و پردازش تصویر می‌شود. مهم‌ترین دلیل اهمیت حذف نویز تصویر این است که به دلیل محدودیت‌های فیزیکی دستگاه‌های مختلف دریافت تصویر و همچنین خطاهای کانال‌های مختلف انتقال بروز نویزهای تصادفی در تصاویر اجتناب ناپذیر است و ممکن است نویز، کیفیت بصری تصویر به دست آمده را به شدت کاهش دهد [۱، ۵، ۶]. در دهه‌های اخیر تولید تصاویر دیجیتالی به صورت چشمگیری افزایش یافته است؛ این تصاویر در زمینه‌های مختلف مانند: پزشکی، سنجش از دور، نظامی و نظارتی و رباتیک تولید

1 Impulse Noise

2 Poisson Noise

3 Additive White Gaussian Noise

4 Probability Density Function

5 Low Level Vision

6 Clean

7 Image Denoising

8 Visual Tracking

9 Image Segmentation

10 Image Classification

11 Medical Imaging

12 Synthetic Aperture Radar

13 Video Surveillance Cameras

14 Ultrasound

15 Magnetic Resonance

16 Computed Tomography

17 Positron Emission Tomography

عنوان توموگرافی یا برش‌نگاری کامپیوتری اشعه ایکس (X-ray CT) شناخته می‌شوند. این تصاویر از ترکیب تعداد بسیار زیادی از تصاویر اشعه ایکس که از زوایای مختلف گرفته شده‌اند، ایجاد می‌شوند. برای تولید تصاویر سی‌تی‌اسکن جسم با استفاده از پرتو اشعه ایکس از زوایای مختلف مشاهده شده و سپس یک تصویر از مقطع داخلی جسم (برش‌های مجازی) به صورت سه‌بعدی بازسازی می‌شود. تصویربرداری سی‌تی‌اسکن از تابش یونیزه استفاده می‌کند و به منظور کاهش اثر تابش یونیزه، تصاویر توموگرافی کامپیوتری با دوز پایین^۵ (LDCT) استخراج می‌شوند. دوز پایین‌تر تابش منجر به تصاویری با کیفیت پایین‌تر می‌شود. دستیابی به تشخیص قابل اعتماد مستلزم آن است که نسبت بین کنتراست بافت مربوطه و دامنه نویز به اندازه کافی بزرگ باشد. پروتکل‌های اسکن با دوز پایین اغلب منجر به افزایش نویز و ایجاد تصاویر تخریب‌شده با مصنوعات رگ‌های غیرایستا می‌شوند. منبع اصلی نویز در تصویر سی‌تی‌اسکن، نویز کوانتومی ناشی از نوسانات آماری کوانتوم‌های اشعه ایکس است. فرض می‌شود نویز تصویر سی‌تی‌اسکن از توزیع پواسون-گوسی مختلط پیروی می‌کند. با این حال، فرآیند بازسازی تصویر سی‌تی‌اسکن از تصاویر اشعه X، تعیین آمار نویز واقعی یک تصویر سی‌تی‌اسکن را دشوار می‌کند؛ از این رو فرض می‌شود که نویز در یک تصویر سی‌تی دارای توزیع آمار گوسی و جمع شونده است [۷، ۱۱-۱۳]

تصاویر توموگرافی گسیل پوزیترون^۶ (PET) یک تکنیک تصویربرداری مولکولی است که تصاویری از جریان خون یا سایر عملکردهای بیوشیمیایی مانند تصاویر متابولیسم گلوکز در مغز ارائه می‌دهد. به بیان دیگر برخلاف سایر تکنیک‌های تصویربرداری که ساختار بدن را نمایش می‌دهند، این تصاویر عملکردهای شیمیایی اندام‌ها و بافت‌ها را در داخل بدن نشان می‌دهند. PET با تشخیص تابش ساطع شده از مواد رادیواکتیو (که به بدن تزریق شده) تصاویری از اندام‌های داخلی بدن تولید می‌کند. در این روش معمولاً از یک اتم رادیواکتیو، مانند کربن-۱۱ که زمان واپاشی کوتاهی دارد استفاده می‌شود. در محل برخورد پوزیترون منتشر شده از ماده رادیواکتیو با الکترون‌های موجود در بافت بدن پرتوهای گامای ساطع می‌شود. دستگاه‌های تصویر برداری PET با تشخیص این پرتوها تصاویر خروجی را تولید می‌کنند. در بسیاری از پژوهش‌ها یک مدل ترکیبی گوسی-پواسون برای توصیف نویز موجود در تصاویر PET در نظر گرفته شده است [۷].

آسیب‌پذیری انواع روش‌های تصویربرداری پزشکی در مقابل انواع مختلف نویز نشان می‌دهد که تکنیک‌های حذف نویز تصویر، یک مرحله حیاتی پیش پردازشی برای بهره برداری از تصاویر پزشکی است. علی‌رغم قابلیت گسترده حذف نویز تصاویر پزشکی، روش‌های حذف نویز موجود، کاستی‌هایی را در مقابله با

امواج صوتی با فرکانس بالا (۲ تا ۱۵ مگاهرتز) با استفاده از یک پروب از دستگاه سونوگرافی به بدن انسان منتقل می‌شوند؛ برخی از این امواج با برخورد به مرز بافت‌های بدن (مثلاً بین مایع و بافت نرم، بافت نرم و استخوان) بازتاب شده و برخی تا رسیدن به مرزهای بعدی و بازتاب شدن به حرکت خود ادامه می‌دهند. امواج منعکس شده توسط پروب دریافت شده و فاصله پروب تا بافت یا اندام (مرزها) با استفاده از سرعت صوت در بافت محاسبه می‌شود. با استفاده از فواصل و شدت بازتاب‌ها یک تصویر دو بعدی را در صفحه نمایش ایجاد می‌شود. تصاویر سونوگرافی به دلیل وجود نویز لکه‌ای^۱ (که ماهیت ضرب شونده دارد) دانه‌ای یا نقطه‌ای به نظر می‌رسند. دلیل ایجاد این نویز لکه‌ای پدیده پراکندگی^۲ است که هنگام انتشار امواج فراصوت از طریق یک محیط بیولوژیکی رخ می‌دهد. این نویز لکه‌ای از نوع ضرب شونده است و وضوح و کنتراست تصویر را به حداقل می‌رساند و لبه‌های تیز تصویر را نیز بدتر می‌کند، بنابراین تشخیص ضایعات کوچک با وجود نویز لکه‌ای دشوار می‌شود. به طور کلی مدل نویز تصاویر سونوگرافی که به طور گسترده در پژوهش‌های مختلف پذیرفته شده است، نویز گوسی ضرب شونده است [۷-۹].

تصاویر رزونانس مغناطیسی (MR): از تصویربرداری MR برای تولید تصاویری از ساختارهای داخلی بدن استفاده می‌شود که با استفاده از سایر روش‌های تصویربرداری پزشکی به خوبی دیده نمی‌شوند. در این تصویربرداری از میدان مغناطیسی و امواج رادیویی استفاده می‌شود. بدن انسان حاوی میلیاردها اتم هیدروژن است. هنگامی که بدن انسان در یک میدان مغناطیسی قرار می‌گیرد اتم‌های هیدروژن با بار مثبت تقریباً به طور یکنواخت در راستای میدان مغناطیسی قرار می‌گیرند. با ارسال پالس‌هایی با فرکانس رادیویی به بدن پروتون‌ها از حالت تراز منحرف می‌شوند، هنگامی که پروتون‌ها به موقعیت اصلی خود باز می‌گردند، انرژی آزاد می‌شود. تصویر MR بر اساس شدت سیگنال دریافتی ایجاد می‌شوند. بزرگی توزیع نویز در سیگنال دریافتی MR برابر با جذر مجموع مربعات دو متغیر گوسی مستقل است، بنابراین نویز در تصویر MR از یک توزیع ثابت رایسی^۳ پیروی می‌کند. در نواحی با نسبت سیگنال به نویز (SNR) بالا در تصویر، توزیع رایسی به گوسی گرایش پیدا می‌کند و در نواحی با SNR پایین، توزیع این نویز به توزیع رایلی^۴ تبدیل می‌شود [۷، ۱۰، ۱۱].

تصاویر سی‌تی‌اسکن یا توموگرافی کامپیوتری (CT) یکی دیگر از انواع روش‌های تصویربرداری پزشکی است که برای تشخیص خونریزی، تومورها، هماتوم و سایر ضایعاتی که ممکن است سکتة مغزی را تکرار کند، استفاده می‌شود. توموگرافی یا برش‌نگاری نوعی تصویربرداری سه‌بعدی است. تصاویر سی‌تی‌اسکن که به

¹ Speckle Noise

² Scattering Phenomenon

³ Rician

⁴ Rayleigh

⁵ Low Dose Computed Tomography

⁶ Positron Emission Tomography

ایفا می‌کند [۱۶, ۱۸]. به همین دلیل تحقیقات گسترده‌ای در خصوص حذف نویز لکه‌ای تصاویر SAR صورت گرفته است. با توجه به ویژگی‌های تصویر و نویز لکه‌ای، روش‌های حذف نویز قبل از تصویربرداری (با استفاده از پردازش چندنمایی^۳) و روش‌های حذف نویز پس از تصویربرداری (مانند استفاده از فیلتر مکانی) برای بهبود تصاویر SAR مورد استفاده قرار گرفته است. در سال‌های اخیر، روش‌های حذف نویز مبتنی بر یادگیری عمیق نیز به روش‌های اصلی حذف نویز تصاویر SAR تبدیل شده‌اند [۱۶].

امروزه دوربین‌های نظارت تصویری به طور گسترده در بسیاری از برنامه‌های کاربردی مدیریت شهر هوشمند و همچنین سیستم‌های کنترل و نظارت خودکار ایمنی برای تضمین ایمنی جاده‌ها مورد استفاده قرار می‌گیرند و می‌توان از داده‌های آن‌ها برای بهبود مسائلی مانند مدیریت ترافیک، کنترل رانندگی و نظارت بر محیط زیست استفاده کرد. بیشتر این برنامه‌ها مبتنی بر الگوریتم‌های تشخیص و ردیابی اشیاء^۴ هستند. با این حال، کیفیت تصویر ویدئویی این دوربین‌ها ممکن است به دلیل تأثیر عوامل خارجی مختلف، به شدت کاهش یافته و مطابق با الزامات چنین الگوریتم‌هایی نباشد. برای مثال انواع شرایط نامساعد آب و هوایی باعث ایجاد نویز در تصاویر می‌شود و تشخیص صحیح اشیاء را دشوار می‌کند. بنابراین، حذف نویز تصاویر طبیعی برای بهره‌برداری از تصاویر دوربین‌های نظارتی نیز بسیار ضروری است [۱۹].

۱-۲- انواع روش‌های حذف نویز تصویر و ویژگی‌ها و چالش‌های آن‌ها

یک روش حذف نویز تصویر باید همزمان با حذف نویز، جزئیات ضروری تصویر را نیز حفظ کند [۱]. به طور کلی یک روش حذف نویز تصویر برای پاسخ به مسائل عملی و کاربردی حذف نویز باید قادر به حذف سطوح مختلف نویز با استفاده از یک مدل واحد باشد. علاوه بر این باید قابلیت‌هایی برای برقراری توازن بین میزان حذف نویز و حفظ جزئیات داشته باشد و بتواند نویزهایی با توزیع مکانی غیر یکنواخت را حذف کند. زمانی که سطح نویز مشخص باشد یا بتوان آن را به خوبی تخمین زد، می‌توان از چنین حذف‌کننده نویزی مستقیماً برای بازیابی تصویر تمیز استفاده کرد. هنگامی که سطح نویز ناشناخته است یا زمانی که برآورد سطح آن دشوار است، این حذف‌کننده نویز به کاربر این امکان را می‌دهد که توازن بین کاهش نویز و حفظ جزئیات را کنترل کند [۲۰].

حذف نویز تصویر یک مساله سنتی است که از مدت‌ها پیش مورد مطالعه محققان قرار گرفته است و در دهه‌های اخیر،

طیف متنوعی از نویزهای ظاهر شده در تصاویر پزشکی نشان می‌دهند.

تصاویر ماهواره‌ای در فعالیت‌های مختلفی مانند نظارت بر محیط زیست، برنامه‌ریزی شهری و مدیریت بحران ضروری هستند. با این حال، محدودیت‌های حسگرهای دریافت تصویر، شرایط نامناسب محیطی و آب و هوایی و خطاهای انتقال اغلب باعث ایجاد نویز در این تصاویر ماهواره‌ای شده و کیفیت تصویر را به شدت کاهش می‌دهند. انواع مختلفی از نویز در تصاویر ماهواره‌ای مشاهده می‌شود. نویز گاوسی که معمولاً ناشی از مشکلات انتقال یا نقص در حسگر است، تصاویر ماهواره‌ای را تحت تأثیر قرار می‌دهد. نویز نمک و فلفل نیز ممکن است به دلیل خطای پیکسل یا خطای انتقال، به صورت پیکسل‌های سفید و سیاه با توزیع تصادفی در این تصاویر ظاهر می‌شود. نویز لکه‌ای یکی دیگر از انواع نویزهای تصاویر ماهواره‌ای است که معمولاً در تصاویر رادار وجود دارد و ناشی از تداخل امواج است. حذف نویز تصویر یکی از پیش نیازهای اساسی بهره‌برداری از تصاویر ماهواره‌ای است و تا به امروز، پژوهش‌های زیادی برای کاهش نویز تصاویر ماهواره‌ای همراه با حفظ جزئیات (بخش‌های با فرکانس بالای) این تصاویر پیشنهاد شده است [۱۴, ۱۵].

رادارهای روزنه مصنوعی^۱ (SAR) بیش از ۳۰ سال است که برای سنجش از دور استفاده می‌شوند. SAR نوعی رادار است که می‌تواند تصاویر دوبعدی یا تصاویر بازسازی شده سه‌بعدی از یک شیء خاص ایجاد کند. SAR یک رادار فعال است که در باند مایکروویو^۲ کار می‌کند. SAR سیگنال‌های مایکروویو جهت‌دار را به سمت زمین ارسال کرده و با دریافت پژواک‌های پراکنده شده از اشیاء سطح زمین، تصاویر دقیق و با وضوح بالایی ایجاد می‌کند. چون این رادار از امواج مایکروویو استفاده می‌کند، امکان تصویربرداری در شب یا شرایط آب‌وهوایی مختلف و همچنین در هنگام وجود ابر یا دود متراکم را فراهم می‌کند. حرکت آنتن رادار می‌تواند وضوح مکانی شیء هدف را بهبود بخشد. از این رو، SAR روی یک سطح متحرک مانند فضاپیما قرار می‌گیرد. مساحتی که دستگاه SAR در واحد زمان پوشش می‌دهد، دیافراگم آنتن مصنوعی را تعریف می‌کند. اگر آنتن بزرگ باشد، دیافراگم فیزیکی است و اگر آنتن متحرک باشد، دیافراگم مصنوعی است. به عبارت دیگر SAR با نمونه‌برداری و تجمیع جابجایی‌های نسبی بین رادار و هدف در فواصل زمانی مختلف، تصاویری با همان وضوح رادار روزنه بزرگ تولید می‌کند. نظریه و فناوری تصویربرداری SAR توجه زیادی را به خود جلب کرده و به سرعت توسعه یافته است [۱۶, ۱۷]. به دلیل ویژگی‌های ذاتی تصویربرداری SAR، این فرآیند اغلب با نویز لکه‌ای همراه است که بر دقت تشخیص هدف تأثیر می‌گذارد. به همین دلیل حذف نویز لکه‌ای نقش مهمی در پردازش تصویر سنجش از دور SAR

³ Multi-View Processing

⁴ Object Recognition and Tracking

¹ Synthetic Aperture Radar

² Microwave

برداری از همبستگی زمانی، پیش‌بینی حرکت صریح در ویدئو است. با این کار، می‌توان فریم‌ها را پیش از فیلتر کردن، نسبت به یکدیگر تراز کرد تا به بازبایی بهتر ویدئو کمک کند. با این حال، از آنجایی که پیش‌بینی حرکت صریح یک مسئله چالش برانگیز و از نظر محاسباتی پرهزینه است، راهکارهای جایگزینی برای روبه‌رو شدن با موضوع حرکت به طور ضمنی ارائه شده است که به صورت بازگشتی ویژگی‌ها را در مراحل زمانی مختلف جمع‌آوری می‌کنند، ولی مدل‌های ارائه شده بر اساس این راهکارها نیز قابلیت اجرای بلادرنگ بر روی سخت افزارهای موجود را ندارند [۲۳].

با استفاده از پیشرفت‌های اخیر در روش‌های مبتنی بر یادگیری عمیق، پژوهش‌های مختلفی در زمینه حذف نویز از ویدئو انجام شده است. در این پژوهش‌ها از تعداد زیادی جفت ویدئوی نویزدار و تمیز برای آموزش شبکه‌های عصبی کانولوشن [۲۴-۲۶] و مبدل‌ها [۲۷]، استفاده شده و موجب پیشرفت چشمگیری در این حوزه شده است. در حالی که مبدل‌ها کیفیت حذف نویز بهتری را نسبت به شبکه‌های عصبی کانولوشن کوچکتر ارائه می‌دهند، اما اجرای آنها نسبت به شبکه‌های کوچکتر کندتر است و تعداد پارامترهای بیشتر آنها منجر به مصرف بالای حافظه هنگام اجرا بر روی تصاویر با وضوح بالا می‌شود (حتی در زمان استنتاج) و سخت‌افزاری را که ممکن است بر روی آن مستقر شوند، محدود می‌کند [۲۲]. روش یادگیری نظارت شده برای حذف نویز تصاویر نویزی واقعی به نتایج خوبی دست یافته است ولی در مقابل، حذف نویز واقعی برای ویدئوهای نویزدار به دلیل چالش جمع‌آوری جفت ویدئوهای تمیز و نویزدار، به خصوص برای صحنه‌های پویا به ندرت مورد مطالعه قرار می‌گیرد. فریم‌های ویدیویی تمیز برای صحنه‌های پویا را نمی‌توان با شاتر با نوردهی طولانی یا با استفاده از میانگین‌گیری چند عکس، (مانند روشی که برای تصاویر ایستا استفاده می‌شود) ضبط کرد (نوردهی طولانی موجب ماتی حرکت در ویدئو می‌شود). به همین دلیل مجموعه‌های داده بزرگ ویدیری شامل جفت ویدئوهای تمیز و تخریب شده برای آموزش شبکه‌های عصبی حذف نویز ویدئو وجود ندارد [۲۴]. با وجود برخی مجموعه داده‌های ویدیویی واقعی، آنها اغلب یا از نظر اندازه محدود هستن و یا از نظر تنظیمات بیش از حد ساده شده‌اند و این موضوع منجر به تعمیم ضعیف شبکه‌های عصبی آموزش دیده می‌شود. برای رفع این مشکل نیز، روش‌های حذف نویز ویدئو بدون نظارت ارائه شده است که با حذف نیاز به جفت ویدئوهای تمیز و نویزی، رویکرد جایگزینی را ارائه می‌دهند. برای مثال در یک رویکرد، مدل‌های آموزش حذف نویز تصویر از پیش آموزش دیده^۱، برای حذف نویز ویدئو تعمیم داده می‌شوند [۲۸]. با وجود پژوهش‌های مختلف در زمینه حذف نویز ویدئو، روش‌های موجود هم از نظر کیفیت حذف نویز برای ویدئوهای واقعی و هم از نظر کارایی محاسباتی

دستاوردهای بزرگی در این زمینه حاصل شده است [۲۱]. دسته‌بندی‌های مختلفی برای روش‌های حذف نویز وجود دارد. بر اساس نوع الگوریتم حذف نویز می‌توان روش‌های حذف نویز را به هفت دسته فیلتر در حوزه مکانی^۱، آستانه‌گذاری (یا فیلتر) تبدیل دامنه^۲، میدان‌های تصادفی^۳، مدل‌های آماری^۴، روش‌های انتشار ناهمسانگرد^۵، روش‌های یادگیری فرهنگ لغت^۶ و روش‌های ترکیبی تقسیم‌بندی کرد [۴]. در یک دسته‌بندی دیگر به طور کلی، روش‌های حذف نویز تصویر را می‌توان به دو دسته، روش‌های مبتنی بر مدل^۷ و روش‌های مبتنی بر یادگیری افتراقی^۸ تقسیم‌بندی کرد [۲۰]. در روش‌های مبتنی بر مدل، مدل‌سازی پیشین تصویر^۹ نقش اصلی را در حذف نویز تصویر ایفا می‌کند ولی در روش‌های مبتنی بر یادگیری، مدل پیشین تصویر با استفاده از روش‌های یادگیری و از مجموعه داده آموزشی شامل زوج تصاویر تخریب شده و تمیز استخراج می‌شود.

با وجود چندین دهه تحقیق در زمینه حذف نویز، روش‌های نوین حذف نویز تصویر هنوز از نظر انعطاف‌پذیری یا کارایی محدود هستند و توانمندی لازم برای بازبایی تصویر تخریب شده با نویزهای دنیای واقعی و در عین حال حفظ ویژگی‌ها و لبه‌های ظریف تصویر را ندارند. علاوه بر این با توجه به اینکه بسیاری از روش‌های حذف نویز مدرن، بر پایه شبکه عصبی هستند و با توجه به انجام پژوهش‌های مختلف در حوزه شبکه‌های عصبی، استفاده از پیشرفت‌های این حوزه ممکن است موجب ارتقای کیفیت روش‌های حذف نویز تصویر نیز شود [۲۰].

حذف نویز از ویدئو نیز یکی دیگر از وظایف اساسی در بینایی ماشین سطح پایین است که همانند حذف نویز تصویر، هدف آن بازبایی ویدئوهای بدون نویز از ویدئوهای تخریب شده است. علاوه بر دلایل تخریب تصویر به دلیل نویز، که منجر به تخریب تصاویر ویدیویی نیز می‌شود، استفاده گسترده از سرویس‌های پخش آنلاین مانند نتفلیکس و دیزنی پلاس و همچنین رسانه‌های اجتماعی مبتنی بر محتوای تولید شده توسط کاربر مانند یوتیوب، اینستاگرام و فیس‌بوک، اهمیت کدگذاری کارآمد ویدئو را افزایش داده است. در کدگذاری ویدئو، الگوریتم حذف نویز برای کاهش آنتروپی فریم‌های ویدیویی و کاهش پهنای باند لازم برای توزیع و دریافت محتوا بسیار ضروری است [۲۲]. تصاویر ویدیویی دارای یک همبستگی قوی در امتداد بُعد زمان هستند، به عبارت دیگر پیکسل‌های موجود در مکان‌های متناظر احتمالاً در فریم‌های متوالی بسیار مشابه هستند. یک استراتژی مناسب برای بهره

1 Spatial Domain Filtering

2 Transform Domain Thresholding

3 Random Fields

4 Statistical Models

5 Anisotropic Diffusion Methods

6 Dictionary Learning Methods

7 Model-Based Methods

8 Discriminative Learning Based

9 Image Prior

۲-۲- روش‌های سنتی حذف نویز مبتنی بر مدل

روش‌های حذف نویز مبتنی بر مدل، از برآوردگر بیشینه احتمال پسین^۶ (MAP) برای پیش‌بینی مقدار تصویر تمیز (x) از نمایش نویزی (y) استفاده می‌کنند [۲۰، ۲۱]. تصویر بازیابی شده با \hat{x} نمایش می‌شود. از دیدگاه بیزی، تخمین بیشینه احتمال پسین x از رابطه (۲) محاسبه می‌شود [۲۱]:

$$\hat{x} = \arg \max_x P(x|y) = \arg \max_x \log P(y|x) + \log P(x) \quad (2)$$

که در آن P(y|x) تابع درست‌نمایی^۷ x و P(x) تصویر پیشین (دانش پیشین تصویر) است. در خصوص نویز گوسی سفید افزودنی (AWGN) با انحراف معیار σ ، تابع هدف را می‌توان به صورت رابطه (۳) بیان کرد [۲۰، ۲۱]:

$$\hat{x} = \arg \min_x \frac{1}{2\sigma^2} \|y - x\|^2 + \lambda R(x) \quad (3)$$

به بخش $\frac{1}{2\sigma^2} \|y - x\|^2$ عبارت صحت داده^۸ گفته می‌شود که تفاوت بین تصویر اصلی و تصویر نویزدار را نشان می‌دهد. $R(x) = -\log P(x)$ یک عبارت منظم‌سازی^۹ و λ ضریب منظم‌سازی است و λ پارامتر منظم‌سازی است [۲۰، ۲۱]. در روش‌های حذف نویز مبتنی بر مدل، نکته کلیدی یافتن یک مدل مناسب پیشین تصویر (R(x)) است. مدل‌های پیشین موفق شامل مدل‌های مبتنی بر گرادیان، مدل‌های خود-مشابه غیرمحل^{۱۰} (NSS)، مدل‌های پراکنده^{۱۱} است [۲۱].

۲-۲-۱- مدل‌های پرکاربرد برای مدل‌سازی دانش پیشین تصویر^{۱۲}

در طول چند دهه گذشته، مدل‌های مختلفی برای مدل‌سازی پیشین تصاویر مورد بهره‌برداری قرار گرفته است. مدل خود-مشابه غیرمحل^{۱۰} (NSS) یکی از متداول‌ترین مدل‌های پیشین مورد استفاده در بینایی کامپیوتر و پردازش تصویر است. هدف مدل‌های خود-مشابه غیرمحل بهره‌برداری از الگوهای تکراری و وصله‌های مشابه موجود در سرتاسر تصویر است [۳۵]. چاترچی و میلانفر [۷] از روش میانگین‌گیری غیرمحل^{۱۰} برای حذف نویز تصویر استفاده کردند و مقدار تخمینی هر پیکسل مانند p را با استفاده از مشابه‌ترین همسایگی p در تمام تصویر تعیین کردند. نمونه‌هایی دیگر از مدل‌های خود-مشابه غیرمحل^{۱۰} (NSS) در [۳۶-۳۸] ارائه شده است.

یکی دیگر از مدل‌های مورد استفاده در بازیابی تصویر مدل‌های پراکنده^{۱۱} است. در مدل‌های پراکنده با تجزیه هر وصله از

برای کاربردهای عملی مناسب نیستند و پژوهش‌های فعال بسیاری در این حوزه در حال انجام است.

ساختار سایر بخش‌های مقاله به این صورت است: در بخش ۲ روش‌های حذف نویز مبتنی بر مدل مورد بررسی قرار می‌گیرند. در بخش ۳ شبکه‌های عصبی حذف نویز بررسی شده و در بخش ۴ معیارهای ارزیابی روش‌های حذف نویز ارائه شده و نتایج پیاده‌سازی روش‌های مختلف حذف نویز با یکدیگر مقایسه شده است و بخش ۵ به نتیجه‌گیری اختصاص دارد.

۲- روش‌های سنتی حذف نویز

۲-۱- روش‌های سنتی حذف نویز مبتنی بر فیلترینگ

از آنجایی که فیلتر کردن یکی از ابزارهای اصلی پردازش تصویر است و در پژوهش‌های پیشین تعداد زیادی از فیلتر فضایی^۱ برای حذف نویز تصویر اعمال شده است که می‌توان آن‌ها را به دو نوع فیلترهای خطی و فیلترهای غیرخطی طبقه‌بندی کرد. با این فرض که طیف فرکانسی نویز، معمولاً از طیف فرکانسی تصویر بیشتر است، فیلترهای فضایی از فیلتر پایین‌گذر بر روی پیکسل‌های تصویر استفاده می‌کنند. فیلترهای فضایی تا حدی قابل قبولی نویز تصویر را حذف می‌کنند اما موجب تار شدن تصویر و از دست رفتن لبه‌های تصویر نیز می‌شوند. در ابتدا، فیلترهای خطی برای حذف نویز در حوزه فضایی به کار گرفته شدند، اما این فیلترها قادر به حفظ بافت‌های تصویر نبودند. برای مثال فیلتر میانگین [۲۹] برای کاهش نویز گاوسی به کار گرفته شده است، ولی این فیلتر ممکن است موجب از بین رفتن بیش از حد جزئیات تصاویر با نویز بالا شود و تصویر را مات کند [۳۰]. برای غلبه بر این نقطه ضعف، فیلتر وینر^۲ [۳۱] استفاده شده است، اما این فیلتر ممکن است لبه‌های تیز تصویر را محو کند. با استفاده از فیلترهای غیرخطی، مانند فیلتر میانه [۲۹، ۳۲] و فیلتر میانه وزن‌دار [۳۳]، می‌توان نویزهای مختلف را، بدون داشتن هیچگونه اطلاعاتی از نوع نویز، از بین برد. فیلتر دوطرفه^۳ [۳۴] یک فیلتر صاف‌کننده^۴ غیرخطی است که لبه‌های تصویر را حفظ می‌کند و به طور گسترده برای حذف نویز تصویر استفاده شده است. در این فیلتر مقدار شدت روشنایی هر پیکسل با میانگین وزنی مقادیر شدت پیکسل‌های مجاور جایگزین می‌شود. این الگوریتم به اجرای جستجوی فراگیر^۵ نیاز دارد و زمانی که شعاع هسته بزرگ باشد، این جستجو هزینه محاسباتی بالایی دارد و کارایی محاسباتی این الگوریتم را با چالش مواجه می‌کند.

⁶ Maximum a Posterior Probability Estimate

⁷ Likelihood Function

⁸ Data Fidelity

⁹ Regularization Term

¹⁰ Nonlocal Self-Similarity Models

¹¹ Sparse Model

¹² Image Prior

¹ Spatial Filters

² Wiener Filtering

³ Bilateral Filtering

⁴ Smoothing

⁵ Brute-Force

پیشین تصویر و نویز (مستقل از تصویر و دارای توزیع یکنواخت گوسی) در نظر می‌گیرند محدودیت‌های حیاتی داشته و کارایی محاسباتی^۴ پایینی دارند. این روش‌ها عموماً شامل یک فرآیند بهینه‌سازی پیچیده در مرحله آزمایش^۵ هستند که فرآیند حذف نویز را زمان‌بر می‌کند. همچنین نمی‌توان از این روش‌ها مستقیماً برای حذف نویزهایی که از نظر مکانی توزیع غیریکنواخت دارند، استفاده کرد. علاوه بر این، روش‌های مبتنی بر مدل معمولاً از مدل‌های پیشین تصویر دست‌ساز^۶ (مانند خود-مشابه غیرمحل و پراکندگی) استفاده می‌کنند که ممکن است برای توصیف ساختارهای پیچیده تصاویر به اندازه کافی قوی نباشند [۲۰]. این روش‌ها شامل پارامترهایی هستند که باید به صورت دستی انتخاب شود. این موضوع امکان انحراف در فرآیند بهبود کارایی حذف نویز به وجود می‌آورد [۴۳].

۲-۳- روش‌های سنتی مبتنی بر یادگیری افتراقی

روش‌های حذف نویز مبتنی بر یادگیری افتراقی به عنوان یک روش جایگزین و برای غلبه بر مشکلات روش‌های مبتنی بر مدل ارائه شدند. هدف از این روش‌ها یادگیری مدل پنهان پیشین تصویر با استفاده از یک مجموعه داده آموزشی شامل جفت تصاویر تخریب شده و واقعی است. در این روش‌ها بهینه‌سازی تکراری در مرحله آزمایش حذف شده است و این موضوع منجر به افزایش سرعت در مرحله استنتاج شده است [۵, ۲۰].

چن و همکاران [۴۴] یک روش به نام مدل انتشار واکنش غیرخطی قابل آموزش^۷ (TNRD) ارائه کردند که در آن با اعمال تعداد ثابتی از مراحل استنتاج گرادیان نزولی^۸، مدل‌سازی پیشین تصویر انجام شده است. انتشار غیرخطی [۴۵, ۴۶] روشی است که در آن هر مرحله انتشار صرفاً شامل عملیات کانولوشن با چند فیلتر خطی است و به همین دلیل کارایی محاسباتی بالایی دارد. اشمیت و راث [۴۷] روش آبشار میدان‌های انقباضی^۹ (CSF) را ارائه کردند. در این روش از مدلی مبتنی بر میدان تصادفی^{۱۰} به عنوان مدل پیشین تصویر و از الگوریتم بهینه‌سازی نیمه درجه دوم باز شده^{۱۱} به عنوان روش یادگیری (برای بهینه‌سازی پارامترها) استفاده شده است. هدف CSF بازیابی تصویر با کیفیت بالا و با سربرار محاسباتی کم است. بر خلاف BM3D، مدل CSF سطح بالایی از موازی سازی را (برای اجرا بر روی GPU) ارائه داده است و کارایی محاسباتی بالایی دارد. دانش پیشین در CSF و TNRD توانایی یادگیری کامل ساختارهای تصاویر پیچیده را

تصویر به ترکیب خطی از چند عنصر پایه، آمار و احتمالات عناصر تصویر کدگذاری می‌شود. به عبارت دیگر هر وصله از تصویر با ترکیب خطی چند فیلتر نمایش داده می‌شود. در این روش عناصر تصویر از یک مجموعه پایه به نام فرهنگ لغت^۱ انتخاب می‌شوند [۳۸]. نمونه‌هایی از مدل‌های پراکنده در [۳۹-۴۱] ارائه شده است.

الاد و اهارون [۴۱] یک روش مبتنی بر مدل پراکنده برای حذف نویز تصویر به نام K-SVD ارائه کردند که از یک فرهنگ لغت بزرگ استفاده کرده که با استفاده از تصاویر تخریب شده و تمیز آموزش دیده است. فرآیند حذف نویز تصویر شامل تفکیک تصویر به وصله‌های همپوشان و سپس اعمال فیلترهایی است که منجر نمایش پراکنده تصویر شود. تصویر حذف نویز شده از میانگین مقادیر وصله‌های همپوشان به دست می‌آید.

دابوف و همکاران [۳۹] با گروه‌بندی وصله‌های مشابه دو بعدی تصویر در قالب آرایه‌های سه بعدی و اعمال فیلترهای سه بعدی به نام فیلترهای مشارکتی^۲، یک الگوریتم قوی به نام BM3D [۳۹] برای حذف نویز گوسی از تصاویر ارائه کردند. در این روش از ترکیب مدل خود-مشابه غیرمحل (NSS) و مدل پراکنده استفاده شده است. حذف نویز در BM3D با اعمال تبدیل بر روی این آرایه سه بعدی وصله‌ها و اعمال کردن قید پراکندگی بر روی خروجی تبدیل و سپس تبدیل بلوک‌های پراکنده به حوزه تصویر انجام می‌شود. برای محاسبه تصویر بازیابی شده وصله‌ها به محل اصلی در تصویر بازگردانده شده و برای وصله‌های همپوشان از میانگین مقادیر وصله‌ها استفاده می‌شود. این روش شامل یک فرآیند تطبیق بلوک است، که برای محاسبات موازی در پردازنده‌های گرافیکی مناسب نیست، بنابراین افزایش سرعت این الگوریتم با بکارگیری معماری‌های موازی امکان پذیر نیست. ژانگ و همکاران [۴۲] نیز یک مدل به نام کمیته‌سازی نرم هسته‌ای وزن دار ماتریس^۳ (WNNM) برای حذف نویز تصویر ارائه دادند که از شباهت‌های غیر محلی و پراکندگی استفاده می‌کند. این روش یک تصویر دریافتی را به وصله‌های کاملاً همپوشانی تجزیه می‌کند و وصله‌های مشابه را گروه‌بندی کرده و آنها را به عنوان ستون‌های یک ماتریس سازمان‌دهی می‌کند. حذف نویز وصله‌ها با کوچک کردن رتبه ماتریس‌های ساخته شده و با به حداقل رساندن نرم هسته ای ماتریس انجام شده است. تصویر بازسازی شده با برگرداندن وصله‌ها به مکان اصلی خود و میانگین وصله‌های همپوشان به دست می‌آید.

روش‌های مبتنی بر مدل عمدتاً مبتنی بر چارچوب احتمالات بیزی هستند و به همین دلیل از تفسیرپذیری خوبی برخوردار هستند. این روش‌ها در زمینه حذف سطوح مختلف نویز انعطاف‌پذیر هستند، ولی به دلیل فرض‌هایی که در مورد دانش

⁴ Computational Efficiency

⁵ Test

⁶ Hand-Crafted Image Priors

⁷ Trainable Nonlinear Reaction Diffusion Model

⁸ Gradient Descent

⁹ Cascade of Shrinkage Fields

¹⁰ Random Field-Based Model

¹¹ The Unrolled Half-Quadratic Optimization Algorithm

¹ Dictionary

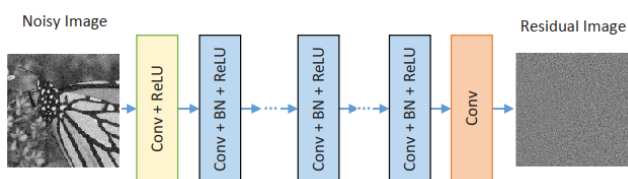
² Collaborative Filter

³ Weighted Nuclear Norm Minimization

۳-۱- افزایش میدان پذیرش^۶ شبکه‌های عصبی کانولوشن حذف نویز

در فرآیند حذف نویز تصویر، اطلاعات زمینه برای بازسازی پیکسل‌های تخریب شده نقش بسیار مهمی را ایفا می‌کند. بزرگ کردن اندازه میدان پذیرش یک راه متداول برای دریافت بیشتر اطلاعات زمینه در شبکه‌های عصبی کانولوشن است. به طور کلی دو راه برای بزرگ کردن میدان پذیرش در CNN وجود دارد، یکی افزایش عمق و عرض شبکه‌های عصبی و دومی استفاده از کانولوشن توسعه یافته^۷ [۵۴].

ژانگ و همکاران [۵] یک شبکه عصبی کانولوشن عمیق با نام DnCNN برای حذف نویز تصویر ارائه کردند و نشان دادند که ترکیب یادگیری رسوبی^۸ و نرمال سازی دسته‌ای^۹ برای حذف نویز تصاویر بسیار موثر است. معماری این شبکه عصبی در شکل (۱) نشان داده شده است. این شبکه از توالی بلوک‌هایی شامل لایه‌های کانولوشن، نرمال سازی دسته‌ای و فعال ساز ReLU تشکیل شده است. بلوک ابتدایی شبکه فاقد نرمال سازی دسته‌ای است و بلوک انتهایی که وظیفه بازسازی تصویر تمیز را بر عهده دارد فقط شامل یک لایه کانولوشن (بدون نرمال سازی دسته‌ای و فعال ساز) است. این شبکه عصبی علی‌رغم معماری ساده، نتایج خوبی در حذف نویز گوسی ارائه کرده است ولی برای هر سطح نویز خاص باید یک مدل خاص از DnCNN آموزش داده شود. علاوه بر این، یک شبکه کلی (DnCNN-B) برای حذف نویز کور ارائه شده است و این شبکه در حذف نویز گوسی با سطوح خاص (مقادیر خاکستری بین ۰ تا ۵۵)، موفق بوده است. علی‌رغم موفقیت‌های قابل توجه این شبکه، DnCNN قابلیت تعمیم پذیری خوبی برای حذف نویز از تصاویر نویزدار واقعی نشان نداده است [۲۰]. علاوه بر وظیفه حذف نویز تصویر، مدل‌هایی از DnCNN برای مسائل وضوح فوق العاده تصویر^{۱۰} و حذف مصنوعات حاصل از فشرده سازی JPG تصویر نیز ارائه شده است.



شکل (۱) معماری شبکه عصبی حذف نویز DnCNN [۵]

با افزایش عمق، آموزش شبکه‌های عصبی مشکل تر می‌شود و افزایش عرض شبکه ممکن است پارامترهای شبکه را به شدت افزایش داده و پیچیدگی مدل حذف نویز را افزایش دهد. برای پاسخ به این مشکلات، کانولوشن‌های توسعه یافته ارائه شدند

ندارد. علاوه بر این، در مرحله آموزش این روش‌ها تعداد زیادی از پارامترهای دست‌ساز دخیل هستند که تعیین این پارامترها، فرآیند آموزش را برای دستیابی به دقت بالا دشوار کرده است. یکی دیگر از اشکالات این روش‌ها این است که برای هر سطح نویز خاص یک مدل خاص را آموزش می‌دهند و در حذف نویز کور^۱ تصویر با محدودیت روبه‌رو هستند [۵].

راث و بلک [۴۸] روشی را به نام Field-of-Experts (FoE) برای یادگیری مدل‌های پیشین تصویر بر اساس میدان تصادفی مارکوف (MRF) و با بهره‌برداری از ایده‌های کلدگاری تصویر پراکنده ارائه دادند. FoE یک مدل پیشین عمومی است که چندین منظم‌ساز (به نام "experts") را با یکدیگر ترکیب کرده است. چارچوب FoE روشی برای یادگیری مجموعه‌ای از MRFها (شامل تمامی فیلترها و همچنین ضرایب ترکیب آن‌ها) از نمونه‌های یک مجموعه داده ارائه داده است. در این روش نیز مانند CSF و TNRD مدل‌سازی پیشین تصویر به صورت دستی انجام شده ولی پارامترهای مدل بر اساس نمونه‌های مجموعه‌ای از تصاویر آموزش داده شده‌اند.

۳-۲- شبکه‌های عصبی حذف نویز تصویر

یکی دیگر از روش‌های رایج حذف نویز تصویر، یادگیری افتراقی ساده است. نمونه‌های این روش‌ها عبارتند از روش‌های مبتنی بر پرسپترون چند لایه^۱ MLP و شبکه‌های عصبی کانولوشن (CNN) [۲۰، ۵]. شبکه‌های عصبی در سال‌های اخیر موفقیت‌های چشمگیری را در زمینه حذف نویز تصاویر ارائه کرده‌اند. مدل‌های مبتنی بر CNN بسیار عمیق، ظرفیت و انعطاف پذیری بسیار بالایی برای بهره‌برداری از ویژگی‌های تصویر ارائه کردند. علاوه بر این در سال‌های اخیر پیشرفت‌های قابل توجهی در روش‌های آموزش شبکه‌های عصبی به دست آمده است که آموزش شبکه‌های عصبی عمیق را بهبود داده است. تعدادی از این موارد عبارتند از: واحد یکسوساز خطی^۳ (ReLU) [۴۹]، نرمال سازی دسته‌ای^۴ [۵۰] و یادگیری رسوبی^۵ [۵۱]. این روش‌ها می‌توانند باعث سرعت بخشیدن به روند آموزش و همچنین بهبود عملکرد حذف نویز در شبکه‌های عصبی عمیق شوند. برگر و همکاران [۵۲]، از MLP برای حذف نویز تصویر استفاده کردند و نتایج قابل رقابت با روش‌های موفق مبتنی بر مدل به دست آوردند. جین و سئونگ [۵۳] یک شبکه عصبی کانولوشن (CNN) برای حذف نویز تصاویر ارائه کردند و ادعا کردند که CNNها قدرت استخراج ویژگی مشابه یا حتی بهتری نسبت به روش‌های موفق مبتنی بر مدل دارند.

⁶ Receptive Field

⁷ Dilated Convolution

⁸ Residual Learning

⁹ Batch Normalization

¹⁰ Single Image Super-Resolution

¹ Blind Image Denoising

² Multi-Layer Perceptron

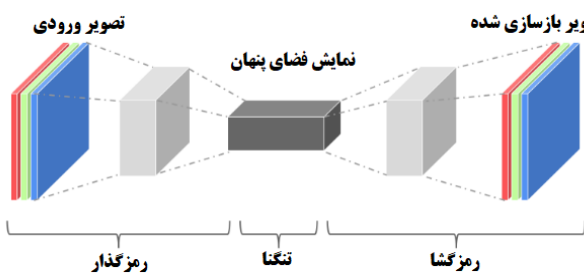
³ Rectifier Linear Unit

⁴ Batch Normalization

⁵ Residual Learning

۳-۲- شبکه‌های عصبی خودرمزگذار حذف نویز^۲

نوع دیگری از شبکه‌های عصبی عمیق که نتایجی برجسته در زمینه حذف نویز به دست آورده است، با نام خودرمزگذارهای حذف نویز [۵۷] شناخته می‌شوند. یک شبکه عصبی خودرمزگذار از دو بخش اصلی تشکیل شده است: بخش رمزگذار^۳ و بخش رمزگشا^۴. بخش رمزگذار با استفاده از تعدادی لایه کانولوشن و لایه‌های نمونه‌برداری کاهشی^۵ یک نمایش پنهان و با ابعاد کوچکتر از داده‌های ورودی را می‌آموزد. بخش رمزگشا با استفاده از لایه‌های کانولوشن و نمونه‌برداری افزایشی تصویر اصلی را از این نمایش پنهان باز تولید می‌کند. از این معماری می‌توان برای انجام هر وظیفه بازسازی تصویر، از جمله حذف نویز تصویر استفاده کرد. معماری کلی این شبکه در شکل (۲) نشان داده شده است [۵۷].



شکل (۲) معماری یک شبکه عصبی خودرمزگذار [۵۷]

زی و همکاران [۵۸] از شبکه خودرمزگذار پراکنده پشته‌ای حذف نویز^۶ (SSDA) برای حذف نویز گوسی استفاده کردند. SSDA نتایج قابل رقابتی با روش‌های موفق حذف نویز مبتنی بر مدل ارائه کرده است. هرچند روش ارائه شده می‌تواند به الگوهای نویز دیده نشده، اما مشابه تعمیم یابد ولی به طور کلی، SSDA می‌تواند فقط الگوهای نویزهایی را حذف کند که در داده‌های آموزشی دیده است.

اتصالات پرشی با پس انتشار^۷ بهتر خطا، باعث می‌شوند آموزش شبکه عصبی سریع‌تر انجام شود. استفاده از این اتصالات موجب بهبود مشکل گرادیان محوشونده^۸ و ساده‌تر شدن آموزش شبکه‌های عمیق شده است. علاوه بر این، این اتصالات جزئیات تصویر را از لایه‌های کانولوشن ابتدای شبکه عصبی به لایه‌های کانولوشن انتهایی آن عبور می‌دهند و موجب بهبود بازیابی جزئیات تصویر می‌شوند [۵۹، ۶۰]. مائو و همکاران [۶۰] یک شبکه رمزگذار-رمزگشای کانولوشن بسیار عمیق با اتصالات پرشی متقارن^۹ به نام RED-Net برای حذف نویز تصویر ارائه کردند

[۵۵]. برای ایجاد توازن بهتر بین دقت و سرعت، ژانگ و همکاران [۵۶] یک شبکه عصبی ۷ لایه را برای حذف نویز، حذف تاری و وضوح فوق‌العاده تصویر تصاویر ارائه کردند که در آن با استفاده از کانولوشن توسعه یافته [۵۴] میدان پذیرش CNN گسترش داده شده است. در سال‌های اخیر شبکه‌های عصبی عمیق حذف نویز و همچنین استفاده از کانولوشن توسعه یافته در کنار کانولوشن معمولی در معماری شبکه‌های عصبی حذف نویز مورد توجه محققان قرار گرفته است و در فصل‌های آینده نمونه‌های آن مورد بررسی قرار خواهد گرفت.

۳-۱-۱- کنترل میزان حذف نویز در شبکه‌های عصبی

کنترل توازن بین میزان حذف نویز تصویر و حفظ جزئیات تصویر، به خصوص در برنامه‌های تعاملی از اهمیت بسیار بالایی برخوردار است. در کاربردهای عملی بازیابی تصویر کاربران تمایل دارند جزئیات بازیابی تصویر خروجی را بر اساس ترجیحات خود کنترل کنند. از این رو روش‌های بازیابی تصویر برای پاسخگویی به سلیقه‌های مختلف کاربران، باید از انعطاف پذیری بالایی در زمینه کنترل توازن بین میزان حذف نویز و حفظ جزئیات تصویر برخوردار باشند. ژانگ و همکاران [۲۰] یک شبکه عصبی کانولوشن سریع و منعطف به نام FFDNet، برای حذف نویز تصویر ارائه کردند که ساختاری مشابه DnCNN دارد و با دریافت یک نقشه سطح نویز^۱ قابل تنظیم به عنوان ورودی، با استفاده از تنها یک مدل FFDNet منفرد نویزهای گوسی با سطوح مختلف و همچنین نویزهایی با توزیع مکانی غیر یکنواخت را حذف می‌کند. در این شبکه عصبی، نقشه سطح نویز، علاوه بر اینکه یک ابزار راهنما برای توزیع نویز در تصویر ورودی است، نقش کلیدی در کنترل توازن بین میزان کاهش نویز و حفظ جزئیات ایفا می‌کند و موجب انعطاف پذیری و ایجاد قابلیت تعاملی FFDNet شده است. برای افزایش میدان پذیرش در این شبکه عصبی، تصویر ورودی در ابتدا به چهار بخش تقسیم شده و بخش‌های مختلف تصویر به یکدیگر متصل شده و سپس به همراه نقشه سطح نویز به شبکه عصبی اعمال می‌شوند. برای FFDNet نتایج خوبی را برای بازیابی تصاویری که با نویز مصنوعی AWGN به صورت دستی تخریب شدند و همچنین تصاویر نویزدار دنیای واقعی ارائه داده است. چون نقشه سطح نویز باید به صورت دستی توسط تجربه انسانی تنظیم شود، تشخیص میزان نویز یکی از چالش‌های استفاده از این روش است و این موضوع باعث شده تا FFDNet از توانایی لازم برای حذف نویزهای واقعی فاصله داشته باشد.

² Denoising Auto-Encoders

³ Encoder

⁴ Decoder

⁵ Down-Sampling Layers

⁶ Stacked Sparse Denoising Auto-Encoders

⁷ Backpropagation

⁸ Vanishing Gradient

⁹ Symmetric Skip Connection

¹ Noise Level Map

جزئیات تصویر، خروجی لایه‌های شبکه رمزگذار بعد از هر نمونه‌برداری، با استفاده از اتصالات پرشی به ورودی لایه‌های رمزگشا متصل شده است و نقشه‌های ویژگی لایه‌های رمزنگار با عملگر الحاق^۲ به نقشه‌های ویژگی لایه‌های رمزگشا متصل شده است. این اتصالات پرشی به حفظ جزئیات تصویر کمک کرده و مانع از حذف آن در میان لایه‌های شبکه عصبی می‌شود. این معماری در وظایف مختلف پردازش تصویر، از جمله حذف نویز تصویر [۶۳-۶۶] مورد استفاده قرار گرفته است. لیو و همکاران [۶۳] با ترکیب مفهوم پراکندگی موجک^۳ و یادگیری عمیق، یک شبکه عصبی به نام MWCNN برای وظایف حذف نویز، وضوح فوق العاده تصویر و حذف مصنوعات ناشی از فشرده سازی JPG ارائه کردند. MWCNN ساختاری مشابه U-Net دارد. برای افزایش میدان پذیرش MWCNN، لایه‌های نمونه‌برداری کاهشی و افزایشی U-Net با تبدیل موجک گسسته دو بعدی^۴ (2DDWT) و معکوس آن جایگزین شده است. علاوه بر این، برای بهبود هزینه محاسباتی MWCNN به جای استفاده از عملگر الحاق برای ترکیب نقشه‌های ویژگی بخش رمزکننده و رمزگشا، در این شبکه از عملگر جمع استفاده شده است.

۳-۳- بهره‌برداری از شباهت‌های غیر محلی در شبکه‌های عصبی حذف نویز

یک تصویر طبیعی اغلب دارای الگوهای محلی تکراری زیادی است و بنابراین یک وصله تصویر همیشه دارای وصله‌های مشابه زیادی در سراسر تصویر است. مدل‌های خود-مشابه غیر محلی برای بهره‌برداری از این شباهت‌ها ارائه شده‌اند. از طریق ادغام این وصله‌های مشابه، می‌توان الگوهای آن‌ها را که در زیر نویزها پنهان شده است، استخراج کرد [۳۵، ۶۷]. مدل خود-مشابه غیر محلی (NSS) یکی از متداول‌ترین مدل‌های سنتی برای بهره‌برداری از الگوهای تکراری و وصله‌های مشابه موجود در سرتاسر تصویر بوده [۳۵] و در پژوهش‌های مختلفی [۳۶، ۳۸، ۳۹، ۴۳] مورد استفاده قرار گرفته است. شبکه‌های عصبی کانولوشن به صورت ذاتی از ویژگی‌های محلی تصاویر استفاده می‌کنند. بهره‌برداری از الگوهای غیر محلی در شبکه‌های عصبی نیز در سال‌های اخیر مورد توجه محققان قرار گرفته است. لفکیمیاتیس [۶۷] یک شبکه عصبی عمیق را برای حذف نویز تصاویر رنگی و خاکستری ارائه کرد که از طریق گروه‌بندی وصله‌های مشابه تصویر و اعمال فیلترهای اشتراکی قابل یادگیری، از شباهت‌های غیر محلی تصاویر طبیعی، بهره‌برداری می‌کند. شبکه پیشنهادی پردازش ویژگی‌های غیر محلی را با یادگیری افتراقی ترکیب کرده است

که در عین سادگی، نتایج قابل رقابت با روش‌های موفق مبتنی بر مدل حذف نویز ارائه کرده است. ساختار این شبکه در شکل (۳) نشان داده شده است. این شبکه شامل لایه‌های متقارن کانولوشن (در بخش رمزگذار) و دکانولوشن^۱ (در بخش رمزگشا) است و اتصالات پرشی به صورت یک در میان نقشه‌های ویژگی لایه کانولوشن را به نقشه‌های ویژگی‌های دکانولوشن متناظر آن‌ها متصل کرده است. در فرآیند رو به جلو، پاسخ (خروجی) لایه کانولوشن به طور مستقیم به لایه دکانولوشن متناظر ارسال شده و در فرآیند انتشار رو به عقب، مقدار خطا به صورت مستقیم از لایه دکانولوشن به لایه کانولوشن منتشر می‌شود.

نویزهای واقعی دارای توزیع غیر یکنواخت و ناشناخته هستند، به همین دلیل برای حذف نویزهای واقعی، حذف کننده نویز باید بتواند با نویزهایی با سطوح و توزیع ناشناخته مقابله کند.



شکل (۳) معماری شبکه عصبی خود رمزگذار حذف نویز RED-Net [۶۰]

به این فرآیند اصطلاحاً حذف نویز کور گفته می‌شود. گو و همکاران [۶۱] یک شبکه عصبی حذف نویز کور به نام CBDNet برای حذف نویزهای دنیای واقعی ارائه کردند. برخلاف روش‌های سنتی که مدل نویز را ثابت فرض می‌کردند، CBDNet همزمان با فرآیند حذف نویز تصویر، تخمین سطوح نویز تصویر را نیز می‌آموزد و با استفاده از این تخمین‌گر سطح نویز، حذف نویز کور را انجام می‌دهد. برای این منظور یک زیر شبکه اختصاصی با ۵ لایه کانولوشن (بدون لایه ادغام و نرمال‌سازی دسته‌ای)، سطح نویز را برای هر پیکسل تصویر ورودی تخمین می‌زند. این ماژول به سازگاری با توزیع مکانی غیر یکنواخت نویزهای واقعی کمک کرده است. معماری زیر شبکه حذف نویز CBDNet مشابه معماری U-Net [۶۲] است. شبکه عصبی U-Net اولین بار برای بخش بندی تصویر ارائه شده است. این شبکه یک معماری سلسله مراتبی مشابه خود رمزگذارها دارد و از اتصالات پرشی نیز بهره می‌برد. معماری این شبکه در شکل (۴) نشان داده شده است. بخش رمزگذار U-Net از لایه‌های کانولوشن، فعال‌ساز ReLU و ادغام بیشینه^۲ تشکیل شده و در هر گام، ابعاد نقشه‌های ویژگی نصف شده و در عین حال عمق آن‌ها (تعداد نقشه‌های ویژگی) دو برابر شده است. هر گام از لایه رمزگشا شامل عملیات نمونه‌برداری افزایشی و یک لایه کانولوشن و فعال‌ساز ReLU است که ابعاد نقشه‌های ویژگی را دو برابر و عمق نقشه‌های ویژگی را نصف می‌کند. برای حفظ

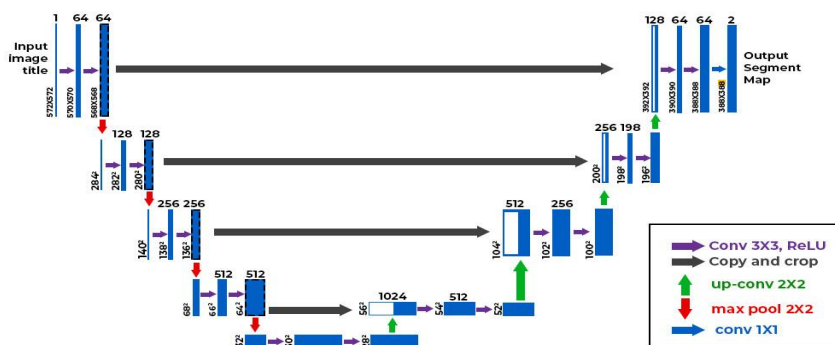
³ Concatenate

⁴ Wavelet Sparsity

⁵ 2D Discrete Wavelet Transform

¹ Deconvolution

² Max Pooling



شکل (۴) معماری شبکه عصبی U-Net [۶۲]

شده و تطبیق بلوک بر روی تصاویر تخریب شده انجام می‌شود. این نوع تطبیق بلوک برای تصاویر به شدت تخریب شده از دقت کافی برخوردار نیست و موجب می‌شود فرآیند تطبیق بلوک به صورت صحیح انجام نشود ولی در [۶۹] مرحله تطبیق بلوک بر روی نقشه‌های ویژگی عمیق که در برابر تخریب مقاوم هستند، اعمال شده است. علاوه بر این فرآیند تطبیق بلوک در [۶۷] به صورت مستقل از شبکه عصبی انجام شده در حالی که ماژول غیرمحلی [۶۹] به طور مشترک با کل شبکه به صورت انتها به انتها آموزش داده شده است. در فرآیند تطبیق بلوک در [۶۷, ۷۰] جستجو در تمام تصویر صورت می‌گیرد ولی در [۶۹] جستجو به صورت کلی نیست و فقط در وصله‌های همسایگی انجام می‌شود.

۳-۳-۱- بهره‌برداری از شباهت‌های غیر محلی با شبکه‌های کانولوشن گرافی (GCN)

با الهام از موفقیت چشمگیر شبکه‌های عصبی کانولوشن عمیق در بینایی کامپیوتری، شبکه‌های کانولوشن گرافی (GCN) برای مدیریت داده‌های ساختاریافته در قالب نمودار، طراحی شده‌اند. برخلاف شبکه‌های عصبی کانولوشن سنتی (CNN) که بر روی ساختارهای داده‌ای شبکه‌ای، مانند تصاویر کار می‌کنند، GCNها یک دسته قدرتمند از مدل‌های یادگیری عمیق هستند که برای کار با داده‌هایی که دامنه نامنظم دارند، طراحی شده‌اند. GCNها برای طیف وسیعی از کاربردها از جمله شبکه‌های اجتماعی، ساختارهای مولکولی و سیستم‌های توصیه‌گر مناسب هستند. به طور خاص GCNها زمانی استفاده می‌شوند که بتوان دامنه داده را به عنوان یک گراف ساختار یافته تعریف کرد و داده‌ها به عنوان بردارهایی بر روی گره‌های این گراف تعریف شوند. یکی از چالش‌های اصلی در طراحی GCNها، تعریف عملیاتی مشابه کانولوشن برای این نوع داده است که ویژگی‌های مانند محلی بودن^۵، ترکیب پذیری^۱ و قابلیت استفاده مجدد از وزن‌ها را که در کانولوشن سنتی وجود دارد، داشته باشد [۷۱].

در پژوهش‌های پیشین عملیات کانولوشن بر روی گراف با دو رویکرد طیفی یا فضایی تعریف شده است. در رویکرد طیفی،

و موجب شده که پیاده‌سازی روش پیشنهادی با استفاده از شبکه عصبی کانولوشن ممکن شود؛ به این ترتیب هم توانایی مدل برای حذف نویز افزایش یافته و هم امکان استفاده از پردازش موازی و GPUهای مدرن، برای یادگیری پارامترهای مدل پیشنهادی فراهم شده است.

پلوت و راث [۶۸] با ارائه شبکه عصبی N3Net، یادگیری عمیق را با مفهوم خود-مشابه غیر محلی ترکیب کردند. برای تطبیق بلوک‌های مشابه در N3Net، یک نوع پیوسته (مشتمل پذیر) از قاعده انتخاب K نزدیک ترین همسایه^۱ (KNN) ارائه شده است و با استفاده از آن یک بلوک جدید غیر محلی، برای ایجاد شبکه عصبی ارائه شده است. معماری N3Net بلوک‌های کانولوشن را با بلوک‌های KNN ترکیب کرده است. بلوک‌های کانولوشن حذف نویز را انجام داده و بلوک‌های KNN نقشه‌های ویژگی را با شکستن آن‌ها به وصله‌های کوچک‌تر و یافتن K نزدیک ترین همسایه برای هر وصله و الحاق وصله‌های مشابه به نقشه ویژگی اصلی، تقویت می‌کنند.

لیو و همکاران [۶۹] یک شبکه عصبی بازگشتی^۲ برای حذف نویز تصویر و وضوح فوق العاده تصویر ارائه کردند که از شباهت‌های غیر محلی تصویر استفاده کرده است. برای بهره‌برداری از شباهت‌های غیر محلی یک ماژول غیر محلی ارائه شده که به طور یکپارچه با شبکه‌های عصبی عمیق موجود ادغام شده و همزمان با آموزش شبکه، بهینه سازی پارامترهای این ماژول نیز انجام شده است. این ماژول فرآیند تطبیق هر بلوک با همسایگی‌ها را با معیار فاصلی اقلیدسی و با اعمال هسته گاوسی جاسازی شده^۳، به صورت خطی انجام می‌دهد. این ماژول می‌تواند با معماری‌های مختلف سازگار شده و امکان بهره‌برداری از شباهت‌های غیر محلی تصویر را برای آن معماری فراهم کند. کیائو و همکاران [۷۰] با ترکیب فرآیند تطبیق بلوک^۴ (استخراج بلوک‌های مشابه) غیر محلی مشابه با [۶۷]، با شبکه TNRD یک شبکه عصبی برای حذف نویز تصویر ارائه کردند. فرآیند تطبیق بلوک در [۶۷, ۷۰] قبل از وارد کردن وصله‌ها به CNN انجام

¹ K-Nearest Neighbors

² Recurrent Neural Network

³ Embedded Gaussian Kernel

⁴ Block Matching

⁵ Locality

⁶ Compositionality

قبل و ۱۶ لایه بعد از لایه غیر محلی وجود دارد.) علاوه بر این ترکیب ویژگی‌های محلی و غیر محلی در [۷۲] با عملگر الحاق انجام شده است.

۳-۴- شبکه‌های عصبی چند شاخه حذف نویز تصویر

نتایج بسیاری از پژوهش‌ها [۵۹, ۷۵-۸۸] نشان داده است که برای بهبود عملکرد مدل‌های مبتنی بر شبکه عصبی به جز افزایش مداوم عمق شبکه عصبی، گسترش عرض شبکه عصبی نیز راهی مفید و کارآمد است. یکی از رویکردهای رایج در طراحی شبکه‌های عصبی که در سال‌های اخیر و در مسائل مختلف، از جمله حذف نویز تصاویر، مورد توجه قرار گرفته است استفاده از ویژگی‌های چند شبکه عصبی در کنار یکدیگر است. شبکه‌های عصبی کانولوشن چند شاخه برای استخراج ویژگی‌های مختلف از تصاویر استفاده می‌شوند. در این شبکه‌ها، هر شاخه از شبکه عصبی برای استخراج ویژگی‌های خاصی از تصاویر آموزش می‌بیند. بکارگیری ویژگی‌های مختلف در کنار یکدیگر، باعث شده تا کیفیت تصاویر بازیابی شده به وسیله شبکه‌های عصبی چند شاخه افزایش یابد. استفاده از شبکه‌های عصبی چند شاخه محدود به وظایف بازیابی تصویر نیست و در سایر وظایف پردازش تصویر [۷۵, ۷۶] نیز کاربرد دارد. تیان و همکاران شبکه عصبی دو شاخه BRDNet [۷۷] را برای حذف نویز تصویر ارائه کردند. این معماری از ترکیب دو شبکه عصبی برای گسترش عرض شبکه استفاده کرده است. معماری شبکه عصبی BRDNet در شکل (۵) نشان داده شده است. در این شبکه علاوه بر معماری دوشاخه، از نرمال‌سازی دسته‌ای، اتصالات پرشی و کانولوشن توسعه یافته و رویکرد یادگیری رسوبی (آموزش نویز تصویر) نیز استفاده شده است. استفاده از معماری دو شاخه باعث شده تا مدل با استفاده از شاخه‌های مختلف، ویژگی‌های مکمل را استخراج کرده و عملکرد حذف نویز بهبود یابد. پان و همکاران [۷۸] یک شبکه عصبی دوشاخه (DualCNN) برای وظایف مختلف بازیابی تصویر، ارائه کردند که شامل دو زیرشبکه با اعماق مختلف است. از زیرشبکه کم عمق برای استخراج ساختارهای تصویر استفاده می‌شود و هدف زیر شبکه عمیق‌تر یادگیری جزئیات تصویر است. بهبود عملکرد شبکه عصبی در BRDNet و DualCNN نشان می‌دهد که استفاده از ساختار چند شاخه CNN برای یادگیری ویژگی‌های مکمل می‌تواند برای حذف نویز تصویر مفید باشد.

هلو و سوسترونک [۷۹] با ارائه یک روش نظری به نام حذف نویز ادغامی^۲ برای پیش‌بینی تصویر تمیز از تصویر نویزدار، یک مدل مبتنی بر شبکه عصبی به نام حذف کننده نویز عمومی و کور ترکیبی^۳ (BUIFD) برای حذف نویز گوسی کور ارائه دادند. BUIFD تنها با یک مدل سطوح مختلف نویز را حذف می‌کند.

عملیات کانولوشن از طریق تبدیل فوریه گراف تعریف شده است. چون در این رویکرد فرض بر این است که گراف ثابت است و مشکل اصلی آن نحوه مدیریت مواردی که ساختار گراف متغیر است. تعریف کانولوشن با رویکرد فضایی برای غلبه بر این مشکلات ارائه شده است. کانولوشن فضایی در گراف، به صورت ترکیب وزنی از مقادیر داده‌های روی گره‌های همسایه (محلی بودن) تعریف شده است [۷۱].

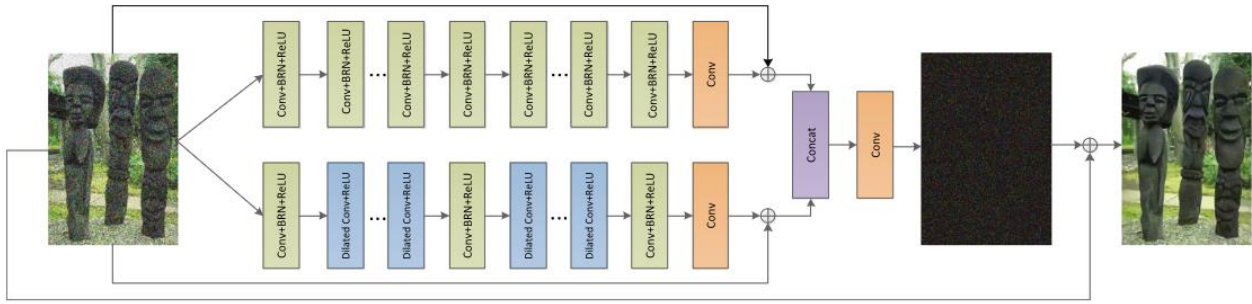
در پژوهش‌های [۷۱-۷۳] از GCN‌ها برای حذف نویز تصویر استفاده شده است. در این پژوهش‌ها از کانولوشن شرطی لبه (ECC) [۷۴] برای طراحی شبکه عصبی گراف استفاده شده است. عملگر ECC با رویکرد فضایی تعریف شده و مقدار آن برابر با میانگین وزن دار گره‌های همسایه است.

ولسیسا و همکاران [۷۱, ۷۳] یک لایه جدید ارائه دادند که از کانولوشن گراف بهره‌برداری می‌کند. این لایه امکان استخراج ویژگی‌هایی را فراهم می‌کند که هم به پیکسل‌هایی که از نظر فضایی مجاور هستند مربوط است و هم به پیکسل‌هایی که از نظر مکانی با هم فاصله دارند ولی در فضای ویژگی به هم نزدیک هستند و همبستگی‌های نهفته دارند. این رویکرد با الگوریتم‌های مبتنی بر تطبیق بلوک عمومی (به عنوان مثال، BM3D) متفاوت است. زیرا این روش مجموعه‌ای از فیلترهای سلسله مراتبی غیرخطی را ایجاد می‌کند که میدان پذیرش غیرمحلی دارند، در حالی که عملیات تطبیق بلوک، مفهومی دست‌ساز از شباهت دارد (مثلاً فاصله اقلیدسی بین وصله‌ها). فیلترهای شبکه عصبی گراف، از شباهت‌های موجود در فضای پنهان ویژگی‌ها ساخته شده و نسبت به معیارهای شباهت دست‌ساز، قدرت بیشتری در استخراج همبستگی‌های موجود در تصویر دارند. در این شبکه عصبی ویژگی‌های محلی با استفاده از کانولوشن معمولی 3×3 استخراج شده و ویژگی‌های غیرمحلی با استفاده از ECC [۷۴] استخراج شده است. برای محاسبه شباهت غیرمحلی (پیکسل‌های همسایه در گراف شباهت)، از الگوریتم K نزدیکترین همسایه بر حسب فاصله اقلیدسی استفاده شده است. گراف شباهت غیر محلی در لایه‌های مختلف شبکه عصبی و بر اساس مقادیر نقشه‌های ویژگی پنهان متناظر با هر پیکسل محاسبه شده است. میانگین ویژگی‌های محلی و غیرمحلی به عنوان نقشه ویژگی خروجی لایه پیشنهادی محاسبه شده است. پژوهش [۷۳] پایه [۷۱] است ولی در [۷۳] با استفاده از روش بهینه‌سازی شیب نزولی مبدایی^۱، معماری شبکه عصبی بهبود یافته و عمق آن به صورت قابل توجهی افزایش یافته است. علاوه بر این راهکارهایی برای رفع مشکل محاسباتی ECC ارائه شده است. بر خلاف [۷۱, ۷۳] که گراف شباهت غیر محلی در لایه‌های مختلف محاسبه می‌شود، در [۷۲] گراف شباهت فقط یکبار و بر اساس ویژگی‌های پنهان لایه‌های میانی شبکه ایجاد شده است. (۱۶ لایه

² Fusion Denoising

³ Blind Universal Image Fusion Denoiser

¹ Proximal Gradient Scent Optimization



شکل (۵) معماری شبکه عصبی دو شاخه BRDNet [۷۷]

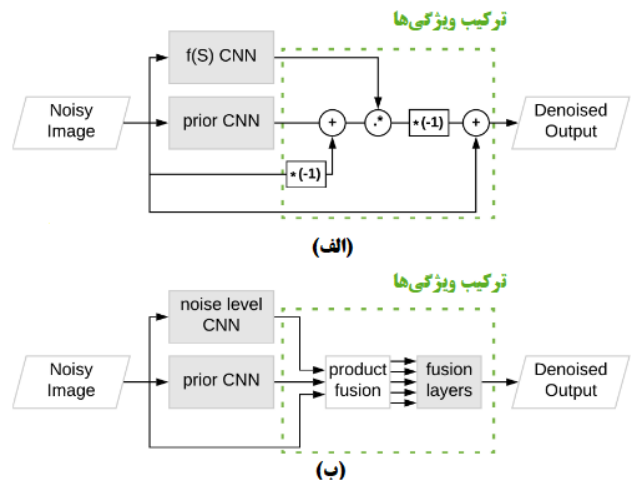
به همراه ضرایبی از آن‌ها به عنوان ورودی شبکه عصبی ترکیب ویژگی اعمال شده است.

کیم و همکاران [۸۰] با معرفی بلوک رسوبی دروازه‌ای^۲ یک شبکه عصبی عمیق دو شاخه به نام ATDNet برای حذف نویز کور تصویر ارائه کردند ATDNet از یک شبکه عصبی عمیق حذف نویز (شاخه اصلی) و یک شبکه عصبی تولید وزن دروازه (شاخه فرعی) تشکیل شده است و تصویر نویزدار و یک نقشه سطح نویز را به عنوان ورودی دریافت کرده و تصویر تمیز را به عنوان خروجی تولید می‌کند. شاخه اصلی این شبکه از بلوک‌های رسوبی دروازه‌ای تشکیل شده است و تصویر نویزدار را به عنوان ورودی دریافت می‌کند. شبکه عصبی تولید وزن دروازه، یک شبکه عصبی ۴ لایه است که نقشه سطح نویز را به عنوان ورودی دریافت کرده و (متناسب سطح نویز تصویر ورودی) ضرایب مورد نیاز بلوک‌های رسوبی دروازه‌ای را تولید می‌کند. استفاده از بلوک‌های رسوبی دروازه‌ای باعث تطبیق پذیری ATDNet با سطوح مختلف نویز شده است. ساختار بلوک دروازه‌ای در شکل (۷) نشان داده شده است. هر یک از این بلوک‌ها از دو لایه کانولوشن و یک فعال‌ساز ReLU در میان آن‌ها تشکیل شده است. این بلوک‌ها ساختاری مشابه بلوک رسوبی معمولی دارند که یک ساختار دروازه کنترل به آن اضافه شده است. این دروازه کنترل مقادیر نقشه‌های ویژگی هر بلوک را کنترل می‌کند. به عبارت دیگر دروازه کنترل می‌کند که نقشه‌های ویژگی لایه‌های کانولوشن این بلوک، چه مقدار در تولید نقشه ویژگی خروجی بلوک (و در فرآیند حذف نویز) تاثیر داشته باشند. خروجی هر بلوک رسوبی دروازه‌ای از رابطه (۴) محاسبه می‌شود. $F_{i,out}$ خروجی بلوک، $F_{i,0}$ ورودی بلوک، α برای کنترل نرخ بروزسانی نقشه‌های ویژگی (مقدار ثابت ۰/۱)، s تابع سیگموئید، G_{σ} خروجی شبکه عصبی تولید وزن دروازه و \circ عملگر ضرب عنصر به عنصر است.

$$F_{i,out} = F_{i,0} + \alpha * \tanh(F_{i,2}) \circ s(G_{\sigma}) \quad (4)$$

پارامترهای شبکه عصبی تولید وزن دروازه همراه با شبکه عصبی حذف نویز آموزش داده شده تا نقشه‌های ویژگی شاخه اصلی، با استفاده از ضرایب شاخه فرعی و متناسب با سطوح نویز به طور مناسب تنظیم شوند. برای استخراج نقشه سطح نویز از تصویر نویزدار، از یک شبکه عصبی تخمین نویز، کم عمق استفاده

در روش حذف نویز ادغامی، تصویر بازیابی شده با استفاده از یک رابطه ریاضی، بر حسب دانش پیشین تصویر (مقدار مورد انتظار^۱ تصویر)، مقدار تصویر نویزدار و نسبت سیگنال به نویز^۲ (SNR) پیش‌بینی شده است. برای پیاده‌سازی این ایده، از شبکه عصبی دو شاخه استفاده شده است یکی از شاخه‌های این شبکه عصبی برای پیش‌بینی دانش پیشین تصویر و یک شبکه عصبی برای پیش‌بینی میزان نویز (تابعی از SNR) استفاده شده است. معماری این شبکه عصبی را در شکل (۶) نشان داده شده است. معماری مورد استفاده در هر یک از این زیرشاخه‌ها، مشابه شبکه عصبی DnCNN است. برای ترکیب ویژگی‌های این دو شبکه عصبی دو ماژول ترکیب طراحی شده است.



شکل (۶) معماری شبکه عصبی دو شاخه BUFD الف) شبکه حذف نویز گوسی ب) شبکه حذف نویز واقعی [۷۹]

ماژول ترکیب مورد استفاده برای حذف نویز گوسی، مبتنی بر عملیات ریاضی مدل نظری حذف نویز ادغامی است (شکل ۶-۶-الف)) ولی برای حذف نویزهای واقعی، به دلیل این که نویز واقعی را نمی‌توان با دانش پیشین گوسی مدل کرد، برای ماژول ترکیب نیز از یک شبکه عصبی با سه لایه کانولوشن استفاده شده است (شکل ۶-۶-ب)). با توجه به اینکه شبکه عصبی امکان ضرب عنصر به عنصر مقادیر ورودی را ندارد، برای پیاده‌سازی مدل نظری حذف نویز ادغامی، خروجی دو زیرشاخه شبکه عصبی استخراج ویژگی،

¹ Expected Value

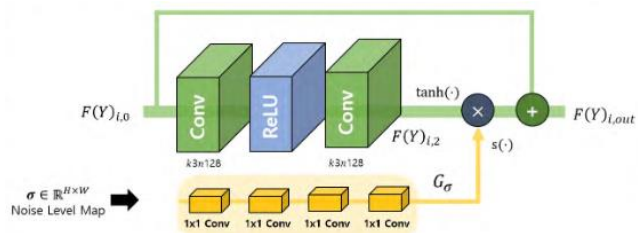
² Signal-to-Noise Ratio

³ Gate-ResBlock

ژو و همکاران [۸۳] شبکه عصبی دوشاخه U-WADN را برای انجام وظایف مختلف بازیابی تصویر ارائه کردند. این شبکه عصبی از دو جزء اصلی تشکیل شده است: شاخه اصلی شبکه که به آن ستون فقرات با عرض تطبیق پذیر^۳ (WAB) گفته می‌شود و شاخه فرعی که انتخابگر عرض^۴ (WS) نام دارد. شاخه WS وظیفه بازیابی میزان هر نوع تخریب در تصاویر ورودی را بر عهده دارد. WAB دارای یک ساختار تودرتو است که چندین شبکه فرعی با عرض‌های مختلف را به صورت یکپارچه در کنار یکدیگر قرار می‌دهد. هدف از این ساختار تقسیم وظایف پیچیده بازیابی تصویر به چند وظیفه ساده و استفاده از شبکه‌های عصبی کوچک‌تر برای انجام این وظایف ساده بازیابی تصویر است. با این طراحی امکان استفاده مجدد از پارامترهای شبکه‌های کوچک‌تر، در شبکه‌های بزرگ‌تر فراهم شده و موجب کاهش هزینه‌های محاسباتی شده است. علاوه بر این، عرض نقشه ویژگی هر یک از زیر شبکه‌های WAB متناسب با نوع تخریب استخراج شده توسط WS تعیین شده و به این صورت هزینه محاسباتی شبکه عصبی به صورت تطبیقی و بر اساس نوع و میزان تخریب تصویر ورودی تعیین شده است.

سو و چو [۸۴] یک حذف کننده نویز مبتنی بر شبکه عصبی کانولوشن (CNN) با نام حذف کننده نویز کور عمومی^۵ (DUBD) ارائه کردند که از مزایای استنتاج بیشینه پسینی^۶ (MAP) و قدرت یادگیری عمیق بهره می‌برد. در طراحی DUBD با استفاده از روش تقسیم و غلبه، مساله حذف نویز تصویر به دو زیر مساله تقسیم شده و برای حل هر مساله از یک شبکه عصبی استفاده شده است. DUBD یک شبکه عصبی دو شاخه است که می‌تواند بدون آگاهی از سطح نویز تصویر، با سطوح مختلف نویز و همچنین نویزهایی با توزیع مکانی غیر یکنواخت نویز مقابله کند. معماری شبکه عصبی DUBD در شکل (۸) نشان داده شده است. شاخه اصلی که عملیات حذف نویز را انجام می‌دهد از تعدادی بلوک تبدیل آفین شرطی^۷ (CATBlock) تشکیل شده است. هر CATBlock از تعدادی بلوک رسوبی شامل دو لایه کانولوشن و تابع فعال‌ساز ReLU در بین آن‌ها، تشکیل شده است. مقادیر نقشه ویژگی آخرین لایه کانولوشن هر CATBlock، با اعمال یک تبدیل آفین (که ضرایب آن در شاخه فرعی تولید می‌شود) اصلاح می‌شود. شاخه فرعی DUBD شبکه تخمین شرایط^۸ (CENet) نام دارد و وظیفه تخمین میزان نویز و تولید ضرایب کنترلی، برای اصلاح نقشه‌های ویژگی بلوک‌های شاخه اصلی را بر عهده دارد. شبکه شرایط دارای ۵ لایه کانولوشن به همراه تابع فعال‌ساز ReLU و

شده است که به صورت مستقل آموزش داده شده است و تصویر نویز دار را به عنوان ورودی دریافت کرده و نقشه سطح نویز را به عنوان خروجی تولید می‌کند.



شکل (۷) معماری یک بلوک رسوبی دروازه‌ای [۸۰]

هی و همکاران [۸۱] با ارائه اتصالات رسوبی قابل کنترل^۱ یک شبکه عصبی عمیق دو شاخه با نام CResMD برای بازیابی تصاویر ارائه کردند که قابلیت کنترل توازن میان میزان بازیابی تصویر و حفظ جزئیات تصویر را دارد و همچنین می‌تواند اثرات تخریب‌های متفاوت (نویز، ماتی تصویر و مصنوعات حاصل از فشرده سازی (JPG) را به صورت جداگانه و تفکیک شده کنترل کند. در این شبکه عصبی هر ماژول وظیفه بازیابی یک نوع خاص از تخریب را بر عهده دارد و اتصالات رسوبی قابل کنترل مشخص می‌کنند هر ماژول به چه میزان در فرآیند بازیابی تصویر تاثیر گذار باشد. یکی از شاخه‌های شبکه عصبی وظیفه استخراج ضریب تاثیر هر ماژول در فرآیند بازیابی تصویر را بر عهده دارد.

لی و همکاران [۸۲] یک شبکه عصبی چند منظوره دو شاخه به نام AirNet برای بازیابی انواع مختلف تخریب تصویر، از جمله حذف نویز، حذف باران، و حذف مه ارائه کردند. این شبکه عصبی با استفاده از یادگیری متقابل^۲ تنها با آموزش یک مدل، عملیات بازیابی تصاویر تخریب شده با خرابی‌های ناشناخته را انجام می‌دهد. یکی از شاخه‌های این شبکه وظیفه استخراج اطلاعات متمایز متناسب با هر نوع تخریب تصویر را بر عهده دارد و شاخه دیگر، با استفاده از این اطلاعات، فرآیند بازیابی تصویر را انجام می‌دهد. یادگیری متقابل برای استخراج اطلاعات تخریب تصویر و ایجاد تمایز بین انواع مختلف تخریب بکار رفته است. این روش علاوه بر پیچیدگی ساختار شبکه، نتایجی ضعیف‌تر از روش‌های مدرن حذف تک وظیفه نویز ارائه داده است.

بر خلاف BRDNet، DualCNN و BUFD که ترکیب ویژگی‌های سطح بالای زیر شاخه‌های مختلف (ویژگی‌های انتهایی شبکه عصبی) موجب بهبود فرآیند بازیابی تصویر شده است در CResMD، ATDNet و AirNet ویژگی‌های استخراج شده توسط یک شاخه به طور مستقیم موجب تغییر در ویژگی‌های شبکه عصبی شاخه دوم می‌شود. در واقع یک شاخه، مقادیر ویژگی‌های شاخه دیگر را کنترل می‌کند به همین خاطر به این شاخه، شبکه عصبی کنترلی نیز گفته می‌شود.

³ Width-Adaptive Backbone

⁴ Width Selector

⁵ Deep Universal Blind Image Denoising

⁶ Maximum a Posteriori

⁷ Conditional Affine Transform Block

⁸ Conditional Estimation Network

¹ Controllable Residual Connections

² Contrastive Learning

فرعی CFSNet تولید شده است استفاده می‌شود. شاخه فرعی CFSNet یک مقدار کنترلی را که بیانگر میزان بازیابی تصویر است، از ورودی دریافت کرده برای هر ماژول اتصالی، ضرایب منحصر به فردی را تولید می‌کند. نقشه‌های ویژگی هر ماژول اتصالی متناسب با این ضرایب تنظیم دقیق می‌شود و عملکرد شبکه متناسب با مقدار ورودی کنترلی اصلاح می‌شود. به این ترتیب CFSNet با دریافت یک ورودی کنترلی قادر به ایجاد توازن بین میزان بازیابی تصویر و حفظ جزئیات تصویر است. هولاً و لی [۸۶] یک معماری جدید و ساده حذف نویز مبتنی بر شبکه عصبی با ترکیب بلوک رسوبی و عملگر تشخیص لبه سوبل ارائه کردند. بخش استخراج ویژگی این شبکه از بلوک‌هایی تشکیل شده که ساختار سه شاخه دارند و هر شاخه از یک یا دو لایه کانولوشن با ابعاد هسته متفاوت تشکیل شده است و ویژگی‌های متفاوتی را استخراج می‌کند. این شبکه در عین سادگی به نتایج قابل رقابتی با سایر روش‌های حذف نویز معروف دست یافته است.

شبکه عصبی چند شاخه فقط برای بهبود عملکرد حذف نویز تصویر استفاده نشده است. یو و همکاران [۸۸] با معرفی بلوک‌های پویا، یک شبکه عصبی چند شاخه برای حذف نویز تصاویر ارائه کردند که باعث کاهش هزینه محاسباتی بازیابی تصویر شده است. با توجه به اینکه پیچیدگی ساختار و الگوهای بخش‌های مختلف تصویر با یکدیگر متفاوت است، بازیابی برخی بخش‌های تصویر ساده‌تر از سایر بخش‌ها است؛ بنابراین هزینه پردازشی مورد نیاز برای بازیابی بخش‌های مختلف تصویر یکسان نیست. این معماری از یک CNN با چندین مسیر مختلف (با هزینه‌های پردازشی مختلف) و یک مسیریاب تشکیل شده است.

مسیریاب با توجه به پیچیدگی ساختار تصویر، مشخص می‌کند هر وصله تصویر با استفاده از کدام مسیر شبکه عصبی بازیابی شود و به این صورت هزینه پردازشی بازیابی تصویر کاهش داده شده است. این معماری توانسته به خوبی با نویزهای واقعی که توزیع مکانی غیریکنواخت دارند مقابله کند.

وو و همکاران [۵۹] یک شبکه عصبی حذف نویز کور کانولوشن دوشاخه با اتصالات پرشی^۹ (DCBDNet) برای حذف نویز گوسی و همچنین نویزهای واقعی تصویر ارائه کردند که در عین سادگی معماری، نتایج قابل رقابتی با سایر شبکه‌های عصبی مدرن ارائه کرده است. برای افزایش انعطاف‌پذیری DCBDNet در برابر انواع مختلف نویز، این شبکه دارای یک زیر شبکه پیش‌بینی نویز ۷ لایه است. برای افزایش میدان پذیرش، تعداد لایه‌های این شبکه نسبت به شبکه‌های پیش‌بینی نویز مشابه بیشتر است. بخش حذف نویز DCBDNet یک شبکه عصبی دو شاخه است؛

یک لایه کانولوشن (بدون فعال‌ساز) است و تصویر نویزی را به عنوان ورودی دریافت کرده و یک نقشه شرایط را به عنوان خروجی تولید می‌کند. برای بهبود کارایی محاسباتی ابعاد این نقشه با درون‌یابی خطی به ابعاد تصویر ورودی تبدیل شده و با اعمال این نقشه ویژگی به یک شبکه عصبی کانولوشن دو لایه با فیلترهای 1×1 (رمزکننده شرایط^۱) ضرایب کنترلی برای تنظیم نقشه‌های ویژگی بلوک‌های شاخه اصلی ایجاد می‌شود. با توجه به اینکه این ضرایب بر اساس میزان نویز تولید شدند، نقشه‌های ویژگی شاخه اصلی شبکه DUBD بر اساس میزان نویز تصویر ورودی (نویزدار) اصلاح شده و به این ترتیب ویژگی‌های استخراج شده توسط شبکه عصبی بهبود می‌یابد.

تیان و همکاران [۸۷] معماری DudeNet را برای حذف نویز تصویر ارائه کردند که شامل بلوک استخراج ویژگی^۲ (FEB)، بلوک بهبود ویژگی^۳ (EB)، بلوک فشرده سازی^۴ (CB) و بلوک بازسازی^۵ (RB) است. FEB با استفاده از دو زیر شبکه عصبی ۱۶ لایه با معماری متفاوت، ویژگی‌های متنوعی را از تصاویر نویزی استخراج می‌کند. یکی از زیر شبکه‌ها با ترکیب کانولوشن معمولی و کانولوشن توسعه یافته^۶ نقاط انرژی بزرگ و کوچک^۷ را با یکدیگر ترکیب کرده تا به طور موثر ویژگی‌های متنوعی را از تصویر نویزی استخراج کند و قدرت شبکه را در پردازش نویزهای واقعی افزایش دهد. زیر شبکه دوم FEB از کانولوشن معمولی، فعال‌ساز ReLU تشکیل شده است. EB ویژگی‌های دو زیر شبکه را از با استفاده از عملگر الحاق با یکدیگر ترکیب کرده و با استفاده از نرمال‌سازی دسته‌ای ناهمگونی توزیع ویژگی‌های دو زیر شبکه را کاهش داده و ویژگی‌های استخراج شده را بهبود می‌بخشد. CB با اعمال کانولوشن 1×1 ویژگی‌های شبکه را فشرده می‌کند تا هزینه‌های محاسباتی را کاهش دهد. RB تصویر تمیز را از ویژگی‌های پردازش شده بازسازی می‌کند.

با رویکردی مشابه DudeNet، وانگ و همکاران [۸۵] با معرفی ماژول اتصالی^۸ شبکه عصبی دو شاخه CFSNet را برای وظایف مختلف بازیابی تصاویر (حذف نویز تصاویر، وضوح فوق العاده تصویر و حذف مصنوعات حاصل از فشرده سازی JPG) ارائه کردند. شاخه اصلی CFSNet، که وظیفه بازیابی تصویر را بر عهده دارد، شبکه پایه نام دارد. شبکه پایه از ماژول‌های اتصالی تشکیل شده است که هر یک ساختاری دوشاخه دارند و نقشه ویژگی یک شاخه به عنوان مقدار مرجع و نقشه ویژگی شاخه دوم به عنوان مقدار اصلاحی نقشه ویژگی شاخه اول استفاده شده است. برای ترکیب این نقشه‌های ویژگی از ضرایبی که توسط شاخه

1 Condition Encoder

2 Feature Extraction Block

3 Enhancement Block

4 Compression Block

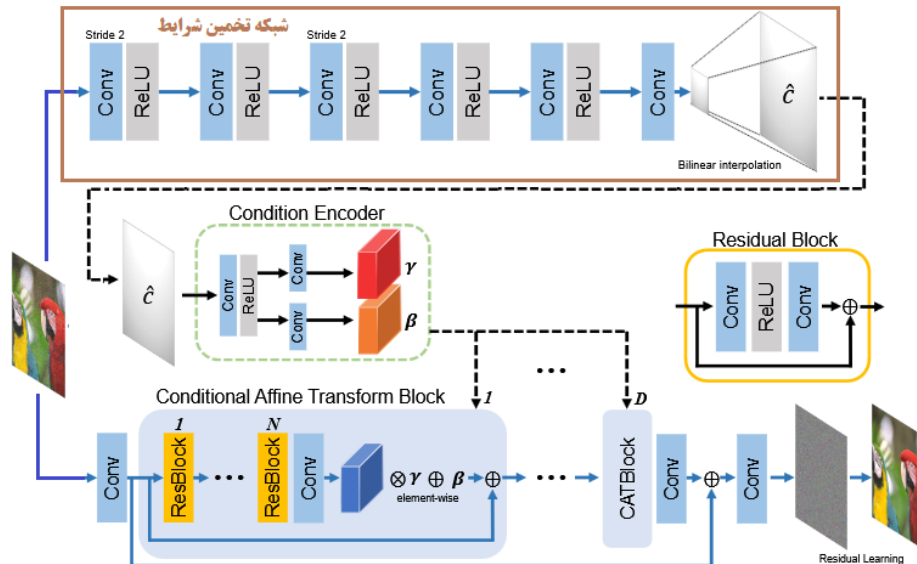
5 Reconstruction Block

6 Dilated Convolution

7 Large and Small Energy Points

8 Coupling Module

⁹ Dual Convolutional Blind Denoising Network With Skip Connection



شکل (۸) معماری شبکه عصبی دو شاخه DUBD [۸۴].

مرتبط را نادیده بگیرد. این مکانیزم می‌تواند به مدل کمک کند تا پیش‌بینی‌های دقیق‌تری داشته باشد و به مدل کارآمدتری تبدیل شود.

در پردازش تصویر نیز استخراج و انتخاب ویژگی‌های مناسب بسیار مهم است. با این حال، به دست آوردن ویژگی‌های مفید برای تصاویر با پس‌زمینه پیچیده آسان نیست [۹۱، ۹۲]. به همین دلیل مکانیزم‌های توجه [۹۳، ۹۴] برای استخراج بهتر ویژگی‌های اطلاعاتی توسعه داده شده و به طور گسترده در بسیاری از زمینه‌های بینایی کامپیوتری [۵۵، ۹۱، ۹۲، ۹۵-۹۷] از جمله حذف نویز تصویر استفاده شده‌اند. مکانیزم توجه می‌تواند ابزاری مفید برای بهبود عملکرد حذف نویز تصویر باشد و تحقیقات بسیاری در این زمینه انجام شده است. در شبکه‌های عصبی حذف نویز از دو دسته از مکانیزم‌های توجه استفاده شده است. در دسته اول از مبدل‌ها [۹۰] برای استخراج ویژگی‌های غیر محلی در شبکه‌های عصبی کانولوشن استفاده شده و در دسته دوم، از مکانیزم‌های توجه مبتنی بر کانولوشن [۹۳، ۹۴] برای بهبود ویژگی‌ها استفاده شده است.

۳-۵-۱ - مبدل‌های حذف نویز

مبدل‌ها [۹۰] که ابتدا برای پردازش زبان طبیعی معرفی شدند، مدل‌هایی هستند که برای مدل‌سازی دنباله داده‌ها طراحی شدند و به دلیل استفاده از مکانیزم توجه برای مدل‌سازی وابستگی‌های دوربرد در دنباله داده‌ها، مورد توجه قرار گرفتند. مبدل‌ها برای غلبه بر چالش‌های موجود در مدل‌هایی مانند شبکه‌های عصبی بازگشتی^۲ (RNN) و شبکه‌های عصبی حافظه طولانی کوتاه مدت^۴ (LSTM) ارائه شدند. مدل‌های سنتی مانند شبکه‌های عصبی بازگشتی (RNN) با مشکل گرادیان محو شونده روبه‌رو

یک شاخه از این شبکه عصبی دارای معماری U شکل است که در آن از نمونه برداری کاهشی و افزایشی برای افزایش میدان پذیرش استفاده شده است. در شاخه دیگر شبکه با استفاده از لایه‌های کانولوشن توسعه یافته میدان پذیرش شبکه افزایش یافته است. اتصالات پرشی در هر دو شاخه DCBDNet برای ترکیب ویژگی‌های لایه‌های کانولوشن مختلف استفاده شده است. ترکیب ویژگی‌های دو زیرشبکه با استفاده از عملگر الحاق انجام شده است.

۳-۵-۲ - استفاده از توجه^۱ در شبکه‌های عصبی حذف نویز تصویر

با توسعه یادگیری عمیق، بسیاری از مدل‌های پیشرفته CNN در وظایف مختلف بینایی ماشین به نتایج بسیار خوبی دست یافتند و شبکه‌های عصبی CNN به روش اصلی و معیار برای حذف نویز تصویر تبدیل شدند. با این وجود، پژوهش‌های مختلف در حوزه CNN‌ها نشان داد که تنها افزایش لایه‌های CNN، نمی‌تواند موجب تطبیق CNN‌ها با انواع مختلف نویز و بهبود عملکرد آن‌ها در پیش‌بینی نویزهای مختلف و با توزیع غیر یکنواخت شود [۸۹]. برای مقابله با این مشکلات، استفاده از مکانیزم توجه برای بهبود ویژگی‌های محلی (که به خوبی به وسیله CNN‌ها استخراج می‌شود) مورد توجه محققان قرار گرفت.

مکانیزم توجه [۹۰] یک تکنیک مورد استفاده در مدل‌های یادگیری عمیق است که به مدل اجازه می‌دهد تا بجای توجه یکسان به همه داده‌های ورودی، به صورت انتخابی بر روی بخش‌های خاصی از داده‌های ورودی تمرکز کند. این تکنیک در هنگام کار با دنباله‌های طولانی داده، مانند پردازش زبان طبیعی یا وظایف بینایی کامپیوتری، بسیار مفید است. این مکانیزم به مدل اجازه می‌دهد بر روی آنچه که مهم است تمرکز کند و بخش‌های کمتر

² Transformers

³ Recurrent Neural Networks

⁴ Long Short Term Memory

¹ Attention

تشکیل شده است. ساختارهای خود توجه و توجه چند وجهی معمولاً بلوک‌های پایه معماری مبدل‌ها را تشکیل می‌دهند. مبدل‌ها با استفاده از خود-توجه مشخص می‌کنند هر بخش از ورودی به چه بخش‌های دیگری وابسته است. این قابلیت به مدل اجازه می‌دهد تا وابستگی‌های دوربرد یا طولانی را شناسایی کند. در معماری مبدل ابتدایی [۹۰]، مکانیزم توجه با استفاده از ضرب نقطه‌ای مقیاس‌پذیر^۶ پیاده‌سازی شده است. در واقع مکانیزم توجه برای ارزیابی شباهت عناصر دنباله ورودی از معیار تشابه کسینوسی استفاده می‌کند. این معیار یک روش خوب برای یافتن شباهت بین دو بردار است. این معیار مقادیری در بازه $[-1, 1]$ تولید می‌کند، که در آن اعداد بزرگتر نشان دهنده شباهت بیشتر و اعداد کوچکتر نشان دهنده عدم شباهت هستند. این مکانیزم توجه با سه ماتریس پرس‌وجو (Q)، کلید (K) و مقدار (V) پیاده‌سازی شده است. هر یک از این ماتریس‌ها یک نمایش کدگذاری^۷ شده از هر یک از کلمات دنباله ورودی را نمایش می‌دهند، به عبارت دیگر هر سطر از این ماتریس‌ها یک بردار است که نشان دهنده یک کلمه در دنباله ورودی است. همانطور که در شکل مشخص است، هر مبدل از سه ماژول توجه تشکیل شده است. ماژول توجه که در بخش رمزگذار قرار دارد، برای محاسبه رابطه بین عناصر دنباله ورودی است. ماژول توجه که در رمزگشا قرار دارد برای محاسبه رابطه بین خروجی‌های رمزگذار (عناصر دنباله خروجی) است. آخرین ماژول توجه، توجه رمزگذار-رمزگشا نام دارد و در انتهای رمزگشا قرار دارد. خروجی‌های توجه رمزگشا به همراه خروجی‌های توجه رمزگذار به این ماژول توجه اعمال می‌شود و برای محاسبه رابطه بین عناصر دنباله خروجی با عناصر دنباله ورودی بکار می‌رود [۹۰، ۹۸].

رابطه (۵) نحوه محاسبه توجه را با استفاده از ضرب نقطه‌ای مقیاس‌پذیر نشان می‌دهد. شکل (۱۰) نمودار ماژول ضرب نقطه‌ای مقیاس‌پذیر را نمایش می‌دهد. از نظر مفهومی همانطور که در این شکل و رابطه (۵) مشخص است، خروجی مکانیزم توجه، ضرایبی از ماتریس V است. این ضرایب مشخص می‌کند که هر کلمه جمله ورودی (Q) چه میزان با هر یک از کلمات این جمله (K) مرتبط است. در پایان برای اینکه وزن‌ها مقادیری بین ۰ و ۱ داشته باشند، تابع SoftMax روی وزن‌ها اعمال می‌شود. d_k ابعاد هر یک از ماتریس‌ها است [۹۰، ۹۸].

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (5)$$

استفاده از چندین توجه موازی متفاوت در قالب ماژول توجه چند-سر برای استخراج حداکثر وابستگی‌ها بین بخش‌های عصبی تسهیل می‌کند و با استخراج ویژگی‌های بهتر از دنباله ورودی، منجر به عملکرد بهتر شبکه عصبی می‌شود.

بودند و این موضوع باعث از دست دادن حافظه بلندمدت می‌شد. به عبارت دیگر RNN‌ها تنها قادر به بهره‌برداری از وابستگی‌های کوتاه‌مدت هستند، در حالی که برای دستیابی به نتایج خوب در پردازش زبان طبیعی^۱ (NLP)، باید از وابستگی‌های بلندمدت نیز بهره‌برداری کرد. علاوه بر این RNN‌ها متن را به صورت خطی پردازش کرده و هر یک از کلمات جمله را تک تک تجزیه و تحلیل می‌کنند و به همین دلیل علاوه بر سرعت آموزش بسیار پایین، قادر به استخراج معنای واقعی جملات نیستند. برای حل این مشکلات، LSTM‌ها ارائه شدند. LSTM‌ها مشکل گرادیان محو شونده RNN‌ها را بهبود بخشیدند و موجب افزایش حافظه مدل در پردازش دنباله‌های داده شدند. استفاده از یک حافظه بلندمدت به نام حالت سلول^۲ در کنار حافظه کوتاه مدت یا حالت پنهان^۳ (که اطلاعات کوتاه مدت از مراحل قبلی در آن ذخیره می‌شود) در LSTM‌ها موجب بهبود عملکرد این شبکه‌های عصبی در پردازش دنباله‌های داده شده است. با این حال، LSTM‌ها هنوز با مشکل پردازش خطی کلمات مواجه هستند. پردازش خطی یا کلمه به کلمه جمله باعث می‌شود که LSTM‌ها نتوانند یک جمله کامل را به طور یکجا تجزیه و تحلیل کنند و این موضوع مانع استخراج معنای واقعی جملات می‌شود [۹۸].

مبدل‌ها برای غلبه بر مشکلات موجود در RNN‌ها و LSTM‌ها مانند عدم بهره‌برداری از وابستگی طولانی ورودی‌ها و پردازش خطی ورودی‌ها (که مانع آموزش موازی شبکه‌های عصبی می‌شود) ارائه شدند. مبدل‌ها کلمات دنباله ورودی را به صورت همزمان پردازش می‌کنند و از مکانیزم خود-توجه برای استخراج وابستگی‌های موجود در کلمات دنباله ورودی استفاده می‌کنند. این موضوع باعث می‌شود که مبدل‌ها بتوانند معنای واقعی متن را بهتر از مدل‌های مشابه استخراج کنند و همچنین امکان استفاده از پردازش موازی و بهره‌گیری از واحدهای پردازش گرافیکی (GPU) را در آموزش این شبکه‌ها فراهم کرده و موجب بهبود سرعت آموزش آن‌ها شده است [۹۸].

ساختار مبدل‌ها در شکل (۹) نشان داده شده است. معماری مبدل‌ها از دو بخش رمزگذار و رمزگشا تشکیل شده است و ساختاری مشابه خودرمزگذارها دارد ولی برخلاف RNN‌ها شامل هیچ شبکه بازگشتی نمی‌شود. هر مبدل از پشت سر هم قرار گرفتن N ماژول رمزگذار/رمزگشا تشکیل شده است. در شکل (۹) در کنار هر ماژول، $\times N$ نشان داده شده است و به معنی کنار هم قرار گرفتن چند ماژول است. رمزگذار، دنباله ورودی را به یک بردار تبدیل می‌کند و رمزگشا این بردار را دوباره به یک دنباله تبدیل می‌کند. هر ماژول رمزگذار/رمزگشا از مکانیزم خود-توجه^۴ و توجه چند-سر^۵ یا چند وجهی به همراه یک شبکه عصبی پیشخور^۱

¹ Natural Language Processing

² Cell State

³ Hidden State

⁴ Self-Attention

⁵ Multi Head Attention

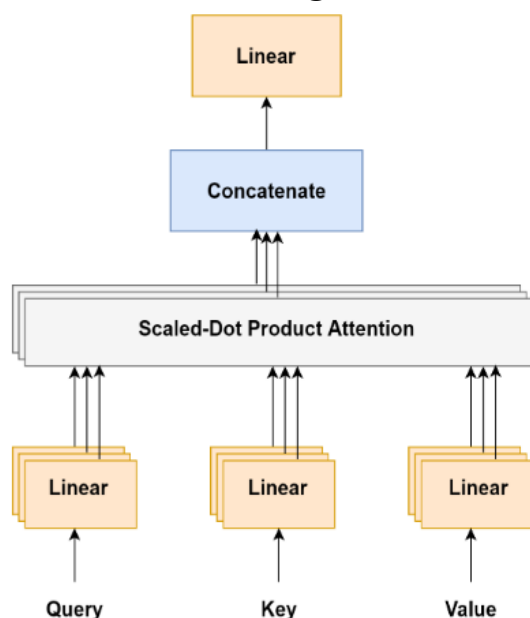
⁶ Feed Forward

⁷ Scaled Dot Product Attention

⁸ Encoded Representation

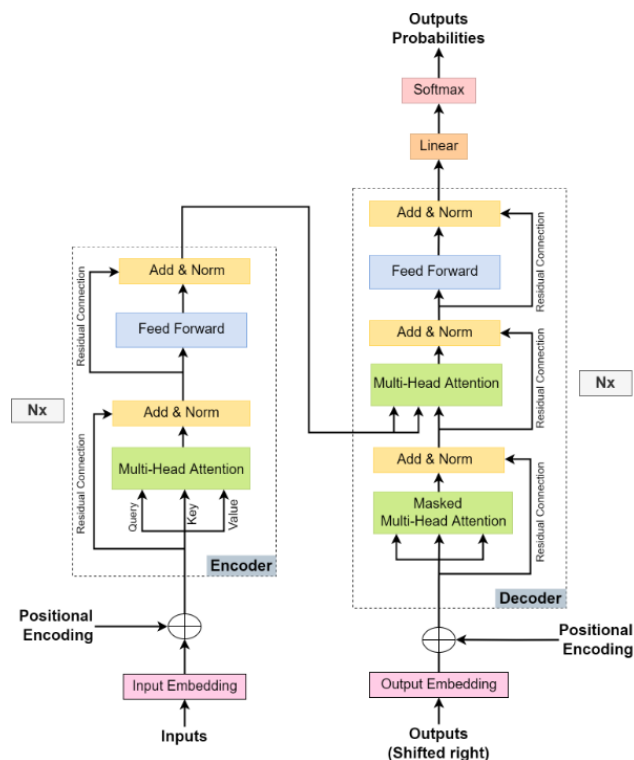
مختلف دنباله ورودی ضروری است. استفاده از توجه چند-سر، یادگیری و ثبت ویژگی‌های متنوع از دنباله ورودی را برای شبکه

مبدل‌ها ابتدا در سال ۲۰۱۷ [۹۰] برای پردازش زبان‌های طبیعی^۱ (NLP) مورد استفاده قرار گرفتند ولی موفقیت فوق‌العاده آن‌ها در حوزه زبانی، محققان را به استفاده از مبدل‌ها در وظایف مختلف دیگر، از جمله پردازش تصویر سوق داد. استفاده از مبدل‌ها در وظایف مختلف پردازش تصویر مانند طبقه‌بندی تصویر [۹۹، ۱۰۰]، تشخیص اشیاء [۱۰۱]، قطعه‌بندی تصویر [۶۶] و مدل‌سازی مشترک بینایی-زبان [۱۰۲] نتایج موفقی را به همراه داشته است. در هنگام استفاده از مبدل‌ها برای وظایف پردازش تصویر، هر تصویر به تعدادی وصله غیر همپوشان تقسیم شده و این وصله‌ها به عنوان دنباله داده ورودی به مبدل اعمال می‌شوند. به این ترتیب مبدل با بررسی وابستگی میان وصله‌ها، وابستگی‌های غیر محلی تصویر را استخراج می‌کند.



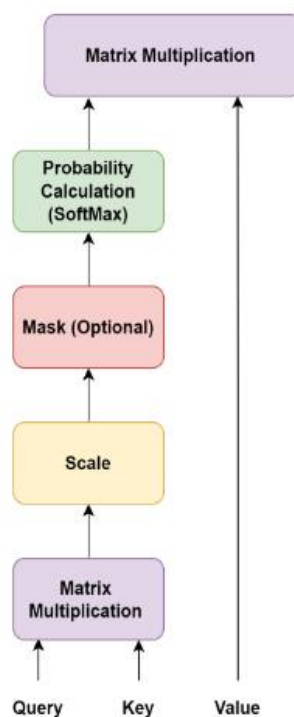
شکل (۱۱) توجه چند-سر یا چند وجهی [۹۸]

مبدل‌های مختلفی برای وظیفه حذف نویز تصویر نیز ارائه شده است. یائو و همکاران [۸۹] یک شبکه عصبی مبتنی بر مبدل به نام DenSformer برای حذف نویز تصویر ارائه کردند که با ترکیب ویژگی‌های محلی و غیر محلی موجب بهبود عملکرد حذف نویز تصویر شده است. لایه‌های کانولوشن ویژگی‌های محلی را استخراج می‌کنند و مبدل‌ها از ویژگی‌های غیر محلی یا سراسری بهره‌برداری می‌کنند؛ در طراحی DenSformer با ترکیب مبدل‌ها و لایه‌های کانولوشن در کنار یکدیگر، ویژگی‌های محلی و سراسری با یکدیگر ترکیب شده و موجب بهبود عملکرد حذف نویز تصویر شده است. این ترکیب ویژگی در معماری DenSformer به وسیله



شکل (۹) ساختار کلی مبدل‌ها [۹۸]

شکل (۱۱) نمودار مکانیزم توجه چند-سر یا چند وجهی را نمایش می‌دهد. در توجه چند وجهی، هر وجه که با k نشان داده شده است، مکانیزم توجه را بر اساس ماتریس‌های وزن قابل یادگیری مخصوص به خود محاسبه می‌کند. خروجی‌های توجه محاسبه‌شده توسط هر وجه به یکدیگر متصل شده و با اعمال یک تبدیل خطی به یک ماتریس واحد تبدیل می‌شوند.



شکل (۱۰) مکانیزم توجه با استفاده از ضرب نقطه‌ای مقیاس پذیر [۹۸]

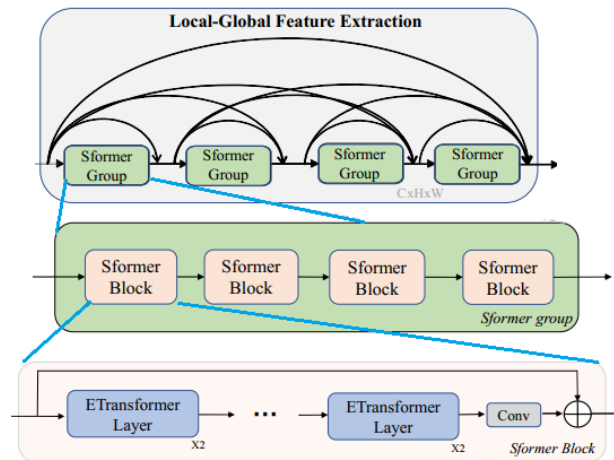
¹ Natural Language Processing

² Global

مکانیزم توجه مبتنی بر کانولوشن استفاده می‌کند. (در بخش بعد جزئیات توجه مبتنی بر کانال ارائه شده است.)
 مبدل پنجره جابجا شده^۲ (Swin) [۱۰۵] یک مبدل تصویر همه منظوره است که هم از مزایای CNNها و هم از مزایای مبدل‌ها بهره می‌برد. Swin به دلیل استفاده از توجه محلی، از مزایای CNNها برای پردازش تصاویر با اندازه بزرگ برخوردار است؛ علاوه بر این با بکارگیری الگوی پنجره جابجا شده از مزایای مبدل‌ها برای مدل‌سازی وابستگی دوربرد بهره می‌برد. Swin با استفاده از ساختار سلسله مراتبی و مکانیسم توجه مبتنی بر پنجره، توازن بین دقت و کارایی محاسباتی ایجاد کرده است. محاسبه خود-توجه^۳ در لایه‌های ابتدایی بر روی وصله‌های کوچک غیر همپوشان انجام شده و با ادغام وصله‌های همسایه در لایه‌های عمیق تر، خود-توجه مجدداً بر روی وصله‌های همپوشان محاسبه شده تا ارتباط بین پنجره‌ای را تضمین کرده و توانایی مدل را برای دریافت ویژگی‌های جهانی^۴ بهبود بخشد. محاسبه خود-توجه در هر پنجره (به جای محاسبه آن در کل تصویر) موجب شده تا Swin پیچیدگی محاسباتی خطی نسبت به ابعاد تصویر داشته باشد.

لیانگ و همکاران [۱۰۶] با استفاده از ماژول‌های مبدل پنجره جابجا شده (Swin) امکان بهره‌برداری از ویژگی‌های غیر محلی را در شبکه عصبی عمیق کانولوشن فراهم کردند و شبکه عصبی عمیق SwinIR را برای وظایف حذف نویز، وضوح فوق العاده تصویر و حذف مصنوعات ناشی از فشرده سازی JPG ارائه کردند. SwinIR از سه بخش استخراج ویژگی‌های کم عمق، استخراج ویژگی عمیق و بازسازی تصویر تشکیل شده است. بخش‌های استخراج ویژگی کم عمق و بازسازی تصویر از یک لایه کانولوشن تشکیل شده است. بخش استخراج ویژگی عمیق بخش اصلی این شبکه عصبی است و از چندین بلوک مبدل Swin رسوبی^۵ (RSTB) تشکیل شده است. هر RSTB دارای چندین لایه مبدل Swin است که در قالب یک ماژول رسوبی در ساختار بلوک قرار گرفته است. این ماژول‌ها برای محاسبه خود توجه محلی در پنجره‌های جابجا شده طراحی شده‌اند، تا به این وسیله از شباهت غیر محلی بهره‌برداری کنند. ژانگ و همکاران [۶۴] با بهره‌گیری از توانمندی لایه کانولوشن رسوبی [۵۱] در استخراج ویژگی‌های محلی، در کنار توانایی مبدل Swin یک شبکه عصبی برای حذف نویز کور تصویر به نام SCUNet ارائه کردند. SCUNet از ساختاری مشابه U-Net [۶۲] به عنوان معماری پایه استفاده کرده که دارای چهار مرحله نمونه برداری کاهشی و افزایشی است و در هر یک از مراحل نمونه برداری از یک بلوک SC^۱ برای ترکیب ویژگی‌های محلی و غیر محلی استفاده

بلوک‌هایی به نام Sformer انجام شده است. هر بلوک Sformer شامل توالی چند مبدل و سپس یک لایه کانولوشن است که ویژگی‌های محلی را با ویژگی‌های سراسری ترکیب می‌کند. بخش اصلی مدل DenSformer، ماژول استخراج ویژگی‌های محلی-سراسری است که از چندین گروه از بلوک‌های Sformer تشکیل شده است. چند بلوک Sformer در کنار یکدیگر یک گروه را تشکیل می‌دهند. برای اتصال گروه‌های Sformer به یکدیگر از اتصالات پرشی متراکم^۱ استفاده شده است تا فرآیند ترکیب ویژگی‌های محلی و غیر محلی را بهبود بخشد و همچنین فرآیند آموزش شبکه را تسهیل نماید. در شکل (۱۲) معماری هر بلوک Sformer و نحوه اتصال گروه‌های Sformer در شبکه عصبی DenSformer نمایش داده شده است.



شکل (۱۲) نحوه قرارگیری ماژول‌های ETransformer در بلوک Sformer و نحوه اتصال گروه‌های Sformer در شبکه عصبی DenSformer [۸۹]

استفاده از معماری متراکم باعث شده تا ترکیبی از ویژگی‌های محلی و سراسری با سطوح مختلف انتزاع (خروجی لایه‌های ابتدایی، میانی و انتهایی شبکه عصبی) در انتهای شبکه عصبی، در اختیار بخش بازسازی تصویر قرار گیرد و این موضوع موجب بهبود بازایی جزئیات تصویر شده است. برخلاف مبدل استاندارد [۹۰] که در آن، از MLP به عنوان شبکه عصبی پیشخور استفاده شده است، در DenSformer از CNN به عنوان شبکه عصبی پیشخور استفاده شده تا با استفاده از لایه‌های کانولوشن، از وابستگی‌های محلی موجود در تصویر نیز استفاده شود. سونگ و همکاران [۱۰۳] نیز با استفاده از مبدل در شبکه عصبی CNN، یک شبکه عصبی حذف نویز به نام DTNet ارائه کردند که در آن از ترکیب ویژگی‌های محلی و سراسری برای حذف نویز استفاده شده است. در DTNet برای بهبود تطبیق‌پذیری شبکه و کاهش هزینه محاسباتی از کانولوشن پویا [۱۰۴] استفاده شده است که عملکردی مشابه مکانیزم توجه مبتنی بر کانال دارد. بنابراین این DTNet در عمل از ترکیب مکانیزم توجه استاندارد (مبدل) و

² Shifted Window Transformer

³ Self-Attention

⁴ Global

⁵ Residual Swin Transformer Blocks

⁶ Swin-Conv

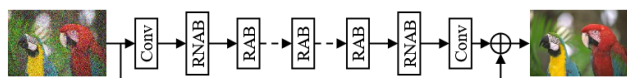
¹ Dense Skip-Connections

فعال‌ساز سیگموئید، ضرایب اهمیت هر کانال را استخراج می‌کند. با ضرب این ضرایب در مقادیر کانال‌های نقشه و ویژگی ورودی، مقدار نقشه‌های ویژگی بر اساس اهمیت آن‌ها اصلاح می‌شود. در ماژول توجه فضایی نیز با فرآیندی مشابه میزان اهمیت هر مکان (هر پیکسل) از تصویر ورودی محاسبه شده و با ضرب نقشه‌های ویژگی در نقشه ضرایب اهمیت، مقدار نقشه‌های ویژگی بر اساس اهمیت مکانی هر پیکسل اصلاح می‌شود.

نسخه‌های مختلفی از مکانیزم توجه ارائه شده است که بر اساس ویژگی‌های مختلف، نقشه‌های ویژگی را مورد تأکید یا سرکوب قرار می‌دهند. دای و همکاران [۱۰۷] با این فرض که تصاویر طبیعی اغلب حاوی اشیاء متنوع و بافت‌های مختلف هستند و این موضوع باعث می‌شود که لایه‌های کانولوشن، نقشه‌های ویژگی با توزیع‌های بسیار پیچیده تولید کنند، یک مکانیزم توجه مبتنی بر کانال ارائه دادند که در آن به جای استفاده از عملگر میانگین و عملگر بیشینه برای تعیین اهمیت کانال، از ترکیبی از ویژگی‌های آماری با مرتبه مختلف استفاده شده است و از این مکانیزم برای بازیابی تصویر استفاده کردند.

دلیل استفاده از ترکیب ویژگی‌های آماری با مرتبه مختلف این است که ویژگی آماری با یک مرتبه نمی‌تواند اطلاعات آماری این ویژگی‌ها را به طور کامل ثبت کند. برای ویژگی آماری مرتبه اول از میانگین، برای ویژگی آماری مرتبه دوم از انحراف معیار استاندارد، برای ویژگی آماری مرتبه سوم از چولگی^۵ و برای ویژگی مرتبه چهارم از کشیدگی^۶ استفاده شده است.

ژانگ و همکاران [۱۰۸] با معرفی بلوک‌های توجه رسوبی محلی و غیرمحلی به عنوان ماژول‌های اولیه طراحی شبکه عصبی بسیار عمیق، یک شبکه عصبی عمیق رسوبی مبتنی بر توجه غیرمحلی^۸ (RNAN) برای وظایف مختلف بازیابی تصویر ارائه دادند. ساختار این شبکه عصبی در شکل (۱۴) نشان داده شده است. این شبکه عصبی یک ساختار متقارن دارد که از توالی یک لایه کانولوشن، یک بلوک توجه رسوبی غیر محلی^۹ (RNAB)، چندین بلوک توجه رسوبی محلی^{۱۰} (RAB) یک بلوک توجه رسوبی غیر محلی (RNAB) و یک لایه کانولوشن تشکیل شده است. ساختار هر بلوک توجه در شکل (۱۵) نشان داده شده است.



شکل (۱۴) ساختار کلی بلوک‌های تشکیل دهنده شبکه عصبی RNAN [۱۰۸]

شده است. هر بلوک SC از یک لایه کانولوشن رسوبی و یک تبدیل Swin تشکیل شده و برای ترکیب ویژگی‌های محلی و غیر محلی از کانولوشن 1×1 استفاده شده است. در پژوهش‌های [۶۵، ۶۶] نیز برای حذف نویز تصویر و بخش بندی تصویر از ترکیب معماری U-Net و مبدل Swin استفاده شده است. در این پژوهش‌ها مبدل Swin به عنوان بلوک اصلی سازنده U-Net (به جای لایه کانولوشن) استفاده شده است در حالی که در [۶۴] یک بلوک جدید برای ترکیب ویژگی‌های محلی و غیر محلی استفاده شده و این بلوک برای بهبود ویژگی‌های استخراج شده با مقیاس مختلف، در ساختار U-Net مورد استفاده قرار گرفته است.

با وجود استفاده از مبدل‌ها در وظایف مختلف بینایی ماشین، وجود برخی مشکلات، کاربرد مبدل‌ها را محدود کرده است. به عنوان مثال، یکی از مشکلات مبدل‌ها استفاده از پیکسل‌های مرزی در تصاویر است، زیرا پیکسل‌های مجاور آن‌ها خارج از محدوده هستند. علاوه بر این مبدل‌ها نمی‌توانند اطلاعات محلی را به خوبی ثبت کنند. بنابراین، نحوه استفاده از مبدل‌ها، به ویژه برای وظایف بینایی ماشین سطح پایین، هنوز با چالش‌هایی روبه رو است [۸۹].

۳-۵-۲ - مکانیزم توجه مبتنی بر کانولوشن در شبکه‌های عصبی حذف نویز

وو و همکاران [۹۳] یکی از مکانیزم‌های توجه معروف در پردازش تصویر را به نام ماژول توجه بلوکی کانولوشن^۱ (CBAM) ارائه دادند. با توجه به اینکه عملیات کانولوشن، ویژگی‌های اطلاعاتی تصویر را با ترکیب اطلاعات فضایی^۲ و کانالی با یکدیگر استخراج می‌کند، CBAM برای تأکید بر ویژگی‌های مهم در امتداد هر یک از دو بعد کانال و بعد مکانی ارائه شده است. معماری CBAM در شکل (۱۳) نشان داده شده است. در این ساختار ماژول توجه فضایی و ماژول توجه مبتنی بر کانال به طور متوالی بر روی ورودی اعمال شده است. در جریان آموزش شبکه عصبی، هر یک از این ماژول‌ها می‌آموزند که به چه بخش از داده (منظور کدام مکان از نقشه ویژگی ورودی در ماژول توجه مکانی و کدام کانال در ماژول توجه مبتنی بر کانال) توجه بیشتری داشته باشد. به این ترتیب CBAM یاد می‌گیرد که کدام اطلاعات را مورد تأکید یا سرکوب قرار دهد و به بهبود جریان اطلاعات در شبکه کمک می‌کند.

ماژول توجه مبتنی بر کانال با اعمال عملیات ادغام میانگین جهانی^۳ (GAP) و ادغام بیشینه جهانی^۴ (GMP) بر روی هر کانال از تصویر ورودی، دو بردار ویژگی را محاسبه کرده و با عبور این مقادیر از یک ساختار تنگنا (شبکه عصبی MLP) و تابع

⁵ Order

⁶ Skewness

⁷ Kurtosis

⁸ Residual Non-Local Attention Networks

⁹ Residual Non-Local Attention Block

¹⁰ Residual Attention Block

¹ Convolutional Block Attention Module

² Spatial-

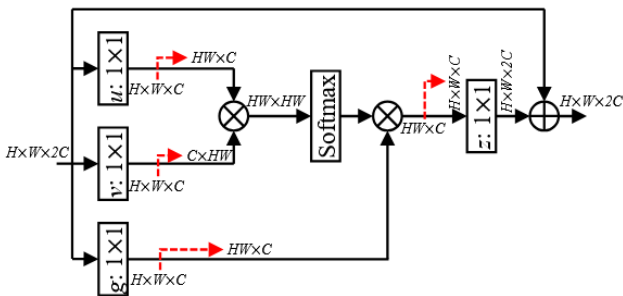
³ Global Average Pooling

⁴ Global Max Pooling

رابطه (۶) عملیات غیر محلی را نشان می‌دهد که در آن i اندیس مکانی ویژگی خروجی را معین می‌کند و j اندیسی است که تمام مکان‌های ممکن را شمارش می‌کند. x و y ورودی و خروجی عملیات غیر محلی هستند. تابع $f(x_i, x_j)$ رابطه بین x_i و x_j را محاسبه می‌کند. تابع $g(x_j)$ نیز نمایشی^۴ از ورودی در موقعیت j را محاسبه می‌کند. برای محاسبه این رابطه می‌توان از تابع softmax استفاده کرد.

$$y_i = (\sum_{j \neq i} f(x_i, x_j) g(x_j)) / \sum_{j \neq i} f(x_i, x_j) \quad (6)$$

همانطور که در شکل (۱۶) نشان داده شده است، از تابع گاوسی توکار^۵ برای محاسبه رابطه جفت عناصر استفاده شده است. برای پیاده سازی این تابع با عملیات ماتریسی از رابطه (۷) استفاده شده است. W_u و W_v ماتریس وزن هستند.



شکل (۱۶) ساختار بلوک‌های غیر محلی موجود در ابتدای شاخه

ماسک شبکه عصبی RNNAN [۱۰۸]

برای محاسبه تابع g نیز از یک عملیات خطی با ماتریس ضرایب W_g استفاده شده است. به این ترتیب میزان رابطه هر جفت مکان از عناصر دنبال ورودی با یکدیگر از طریق رابطه (۷) محاسبه می‌شود.

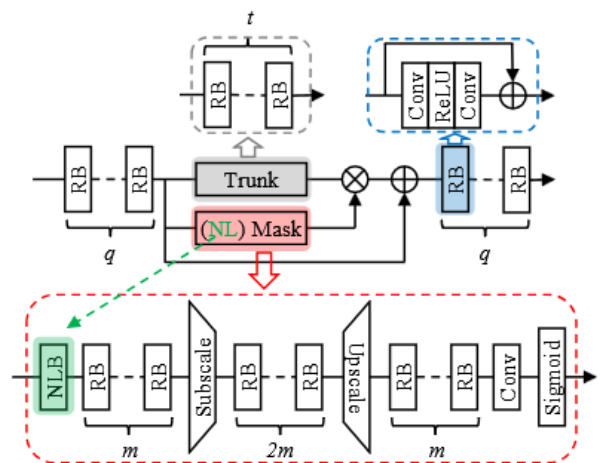
$$f(x_i, x_j) = \frac{\exp(u(x_i)^T v(x_j))}{\exp((W_u x_i)^T (W_v x_j))} \quad (7)$$

بعد از محاسبه رابطه هر جفت عناصر ورودی، می‌توان خروجی لایه غیر محلی را برای هر مکان با اندیس i از رابطه (۸) محاسبه کرد. W_z ماتریس وزن خروجی است.

$$Z_i = W_z y_i + x_i = W_z \text{softmax}((W_u x_i)^T (W_v x_j) g(x_j)) + x_i \quad (8)$$

خروجی لایه ماسک برای وزن دهی به نقشه‌های ویژگی شاخه تنه استفاده شده است و به این وسیله ویژگی‌های مهم شاخه تنه تقویت شده و ویژگی‌های کم ارزش، سرکوب می‌شوند. استفاده از ترکیب توجه محلی و غیر محلی موجب بهبود عملکرد شبکه عصبی در وظایف مختلف بازیابی تصویر مانند حذف نویز، وضوح فوق العاده تصویر^۶ و حذف مصنوعات حاصل از فشرده‌سازی^۷ تصویر شده است.

تنها تفاوت بلوک‌های توجه رسوبی محلی و غیر محلی، در وجود یک بلوک غیر محلی^۱ (NLB) در ابتدای شاخه ماسک است. هر بلوک توجه رسوبی از یک ماژول توجه (شبکه عصبی دوشاخه) و تعدادی بلوک رسوبی (RB) قبل و بعد از این ماژول تشکیل شده است. ساختار هر بلوک رسوبی نیز در شکل (۱۵) نشان داده شده است؛ این بلوک ساختاری ساده دارد و از دو لایه کانولوشن و یک فعال‌ساز ReLU در بین آن‌ها تشکیل شده است. ماژول توجه یک شبکه عصبی دو شاخه است و از دو شاخه تنه^۲ و ماسک تشکیل شده است. شاخه تنه نیز از تعداد t بلوک رسوبی تشکیل شده است. هدف اصلی شاخه ماسک این است که اطلاعات با دامنه بزرگتر (اندازه میدان پذیرش بزرگتر) را استخراج کرده و به این وسیله مکانیزم توجه پیچیده تری را پیاده سازی کند.



شکل (۱۵) معماری یک بلوک توجه رسوبی محلی (و غیر محلی) شبکه عصبی RNNAN. خط چین قرمز رنگ معماری بخش ماسک، خط چین خاکستری بخش تنه و خط چین آبی ساختار بلوک رسوبی است [۱۰۸]

شاخه ماسک یک شبکه عصبی عمیق است که برای پیاده‌سازی آن از لایه‌های کانولوشن و دکانولوشن^۳ با گام بزرگ استفاده شده است تا اندازه میدان پذیرش افزایش یابد و وزن‌های توجه ترکیبی (ترکیب توجه مبتنی بر کانال و توجه فضایی) بر اساس اطلاعات کلی تصویر استخراج شوند. به جای این ساختار می‌توان از عملگر ادغام پیشینه استفاده کرد ولی دلیل استفاده از این ساختار این است که عملگر ادغام پیشینه منجر به از دست رفتن بخش مهمی از جزئیات تصویر می‌شود. علاوه بر این ساختار، برای در نظر گرفتن اطلاعات کلی تصویر، یک عملگر غیر محلی در ابتدای بلوک توجه رسوبی غیر محلی استفاده شده است. ساختار بلوک غیر محلی (NLB) شاخه ماسک در شکل (۱۶) نشان داده شده است. $H \times W \times C$ نشان دهنده تعداد C نقشه ویژگی با ابعاد $H \times W$ است. علامت \oplus نشان دهنده جمع عنصر به عنصر و علامت \otimes نشان دهنده ضرب ماتریسی است.

⁴ Representation

⁵ Embedded Gaussian Function

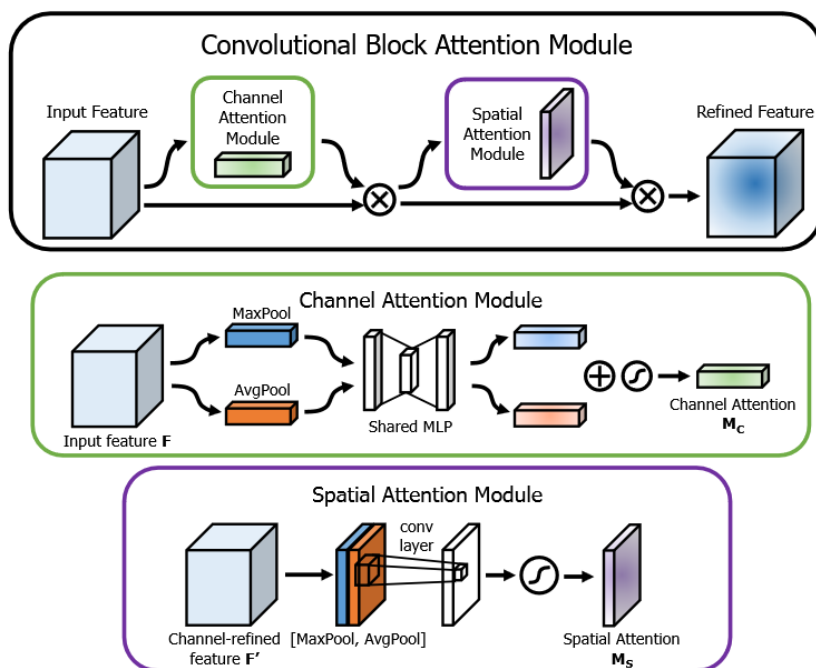
⁶ Image Super-Resolution

⁷ Image Compression Artifacts Reduction

¹ Non-Local Block

² Trunk

³ Deconvolution



شکل (۱۳) معماری ماژول توجه CBAM و معماری ماژول توجه مکانی و ماژول توجه مبتنی بر کانال مورد استفاده در CBAM [۹۳]

که برای استخراج ویژگی‌های اولیه از ورودی نویزی استفاده شده است. خروجی این لایه به ماژول یادگیری ویژگی رسوبی در رسوبی اعمال شده است. این ماژول از ماژول‌های توجه بهبود یافته^۲ (EAM) تشکیل شده است که به صورت آبشاری در کنار هم قرار گرفته‌اند. ویژگی‌های خروجی لایه نهایی به ماژول بازسازی اعمال شده است. این ماژول نیز از یک لایه کانولوشن تشکیل شده است.

معماری رسوبی در رسوبی بکار رفته در RIDNet دارای تعداد زیادی اتصال رسوبی است، که بخش‌های مختلف شبکه را با معماری رسوبی به یکدیگر متصل می‌کند. این اتصالات برای تسهیل جریان اطلاعات با فرکانس پایین استفاده شده تا توانایی شبکه برای بازسازی تصاویر تمیز از ورودی‌های نویز دار را افزایش دهد. RIDNet از چهار ماژول EAM تشکیل شده و هر EAM از تعدادی لایه کانولوشن به همراه فعال ساز ReLU با اتصالات پرشی و یک ماژول توجه تشکیل شده است. معماری EAM در شکل (۱۸) نشان داده شده است. بخش اول EAM، یک بلوک رسوبی است که دارای یک ساختار دو شاخه است و هر شاخه دارای لایه‌های کانولوشن با ضریب توسعه^۳ متفاوت است و اطلاعات مختلفی از ورودی استخراج می‌کند. در هر EAM با اعمال هسته کانولوشن 1×1 برای افزایش سرعت ویژگی‌ها فشرده سازی شده و در نهایت یک ماژول توجه (مبتنی بر کانال)، وزن نقشه‌های ویژگی مهم را افزایش می‌دهد. ماژول توجه در RIDNet مشابه ماژول توجه مبتنی بر کانال CBAM [۹۳] است.

هر یک از بلوک‌های RDAB و RCAB یک بلوک مبتنی بر توجه است که از یک شاخه اصلی (شاخه تنه) و یک شاخه ماسک تشکیل شده است. معماری بلوک‌های RDAB و RCAB در شکل (۱۷) نشان داده شده است. شاخه تنه پردازش ویژگی را انجام داده و از دو بلوک رسوبی تشکیل شده است. شاخه ماسک در RCAB از دو بلوک رسوبی تشکیل شده و در RDAB از یک بلوک توسعه یافته و یک بلوک رسوبی تشکیل شده و برای یادگیری ضرایب اهمیت (وزن دهی) نقشه‌های ویژگی شاخه اصلی استفاده شده است. در این شبکه با استفاده از اتصالات پرشی طولانی، ویژگی‌های کم عمق، به لایه‌های خروجی شبکه پیوند داده شدند تا هم اطلاعات سطح پایین، در لایه‌های انتهایی شبکه در دسترس باشد و هم انتشار رو به عقب گرادیان و پیش‌بینی در سطح پیکسل‌ها تسهیل شود. با توجه به اینکه پشت سرهم قراردادن ساده ماژول‌های توجه منجر به افت عملکرد شبکه می‌شود [۱۰۹]، برای مقابله با این مشکلات، از یادگیری رسوبی مبتنی بر توجه برای آموزش شبکه استفاده شده است. بلوک توجه در RDAB علاوه بر افزایش میدان پذیرش، به دلیل استفاده از کانولوشن 1×1 ، ویژگی‌های محلی و غیر محلی را استخراج می‌کند.

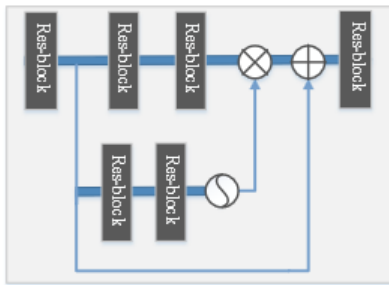
انور و بارنز [۹۷] شبکه حذف نویز تصویر رسوبی (RIDNet) را ارائه کردند که یک مدل حذف نویز کور برای حذف نویز تصاویر تخریب شده با نویزهای واقعی است. RIDNet از سه ماژول اصلی تشکیل شده است: ماژول استخراج ویژگی، ماژول یادگیری ویژگی رسوبی در رسوبی^۱ و ماژول بازسازی تصویر. ماژول استخراج ویژگی یک لایه کانولوشن است

² Enhancement Attention Modules

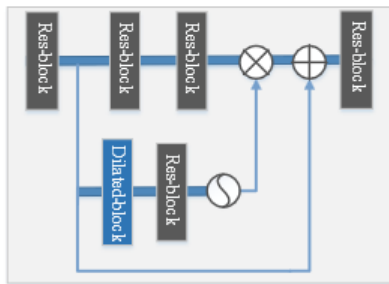
³ Dilations

¹ Feature learning Residual on the Residual

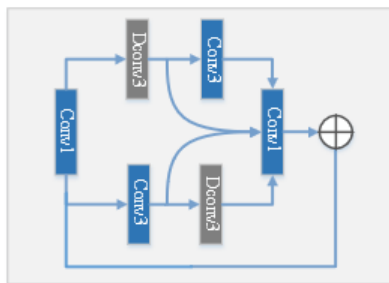
هر دو به عملکرد حذف نویز قابل رقابتی با سایر روش‌های مدرن دست یافتند ولی تنها یک نوع مکانیسم توجه در آن‌ها استفاده شده است و این موضوع ممکن است منجر به نادیده گرفتن سایر روابط متقابل ویژگی‌های استخراج شده توسط لایه‌های کانولوشن شود. ونکونگ و همکاران [۹۲] با الهام از موفقیت ساختار CNN دو شاخه و مکانیسم توجه در شبکه‌های عصبی حذف نویز تصویر، یک شبکه عصبی مبتنی بر توجه رسوبی دو شاخه^۷ (DRANet) برای حذف نویز کور تصویر ارائه کردند، که هم از معماری مدل عریض و هم از مکانیسم توجه استفاده می‌کند.



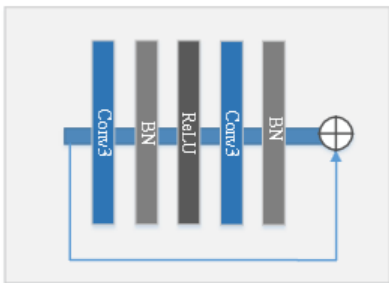
(الف)



(ب)



(ج)



(د)

شکل (۱۷) بلوک‌های تشکیل دهنده شبکه عصبی RDAN. (الف) معماری یک بلوک توجه کانولوشن رسوبی (RCAB) (ب) معماری یک بلوک توجه توسعه یافته رسوبی (RCAB) (ج) معماری بلوک Dilated-block (د) معماری بلوک Res-block.

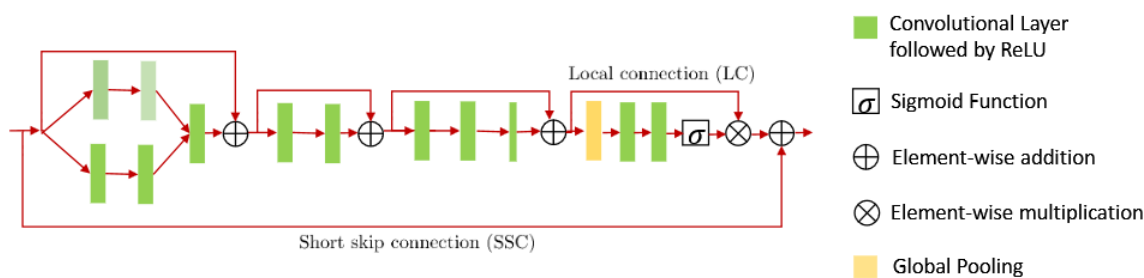
تیان و همکاران شبکه عصبی ADNet [۵۵] را ارائه کردند که در آن، مکانیزم توجه موجب بهبود فرآیند استخراج اطلاعات نویز در تصاویر نویزدار دارای زمینه پیچیده، شده است. مکانیزم توجه استفاده شده در این مقاله نیز توجه مبتنی بر کانال^۱ هست. معماری شبکه ADNet از چهار بلوک تشکیل شده است: بلوک خلوت^۲ (SB)، بلوک بهبود ویژگی^۳ (FEB)، بلوک توجه^۴ (AB) و بلوک بازسازی تصویر^۵ (RB). معماری این شبکه در شکل (۱۹) نمایش داده شده است. بلوک SB با استفاده از کانولوشن‌های معمولی و توسعه یافته ویژگی‌های مفید را از تصویر نویزدار استخراج می‌کند. توالی این دو نوع لایه کانولوشن در کنار یکدیگر برای افزایش میدان پذیرش شبکه عصبی و همچنین بهبود عملکرد حذف نویز آن است. این ترتیب قرارگیری لایه‌های کانولوشن معمولی و توسعه یافته برای بهبود ویژگی استخراج شده و همچنین حذف پدیده شبکه‌ای^۶، که در اثر استفاده از لایه‌های کانولوشن توسعه یافته پدید می‌آید، بکارگرفته شده است. استفاده از کانولوشن توسعه یافته در، این بلوک موجب کاهش عمق شبکه شده و عملکرد حذف نویز و کارایی محاسباتی را بهبود داده است. FEB برای بهبود ویژگی‌های مدل حذف نویز است. FEB یک شبکه عصبی چهار لایه است که ویژگی‌های سراسری و محلی را با یکدیگر ترکیب کرده و ویژگی‌های بهتری را به عنوان خروجی تولید می‌کند. چون در شبکه‌های عمیق ممکن است جزئیات تصویر ورودی، در لایه‌های کانولوشن از بین برود، این بلوک تصویر نویزدار را نیز به عنوان ورودی دریافت می‌کند و با ترکیب آن با ویژگی‌های استخراج شده توسط SB موجب بهبود ویژگی‌های مورد استفاده در حذف نویز تصویر می‌شود. منظور از ویژگی‌های سراسری، تصویر نویزی ورودی است و خروجی بلاک SB به عنوان ویژگی‌های محلی در نظر گرفته می‌شود. بلوک AB با استفاده از یک لایه کانولوشن 1×1 ویژگی‌های به دست آمده را در قالب یک بردار فشرده کرده و از این بردار به عنوان وزن‌های توجه مبتنی بر کانال استفاده کرده و ویژگی‌های بهتر و برجسته‌تر نویز را از زمینه پیچیده تصویر استخراج می‌کند.

چون ورودی AB ترکیبی از ویژگی‌های محلی و سراسری است، به خوبی می‌تواند الگوهای نویز را سرکوب کرده و الگوهای بافت تصویر را تقویت کند. همانطور که در شکل (۱۹) مشخص است این شبکه از رویکرد یادگیری رسوبی استفاده کرده و نویز تصویر را پیش‌بینی می‌کند.

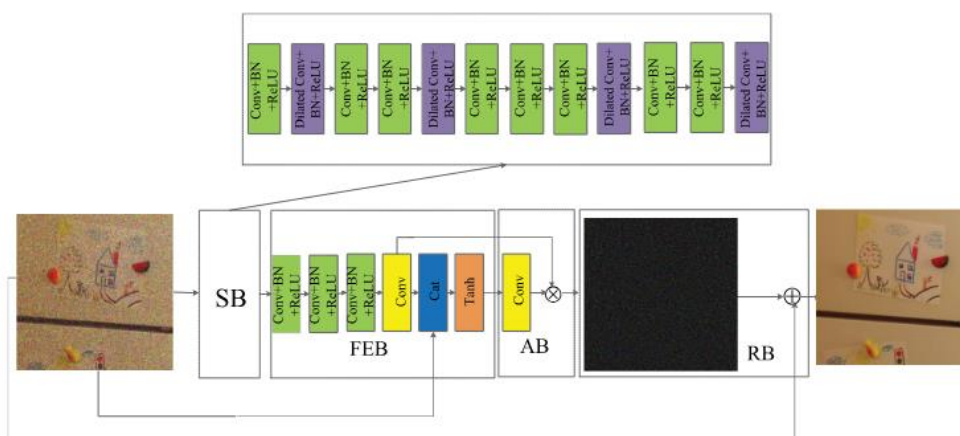
RIDNet با استفاده از ماژول توجه برای ویژگی‌های با ارزش اهمیت قائل می‌شود و ADNet اطلاعات نویز را از طریق بلوک‌های توجه استخراج می‌کند. اگرچه ADNet و RIDNet

1 Channel Attention
2 Sparse Block
3 Feature Enhancement Block
4 Attention Block
5 Reconstruction Block
6 Gridding Phenomenon

⁷ Dual-Branch Residual Attention Network



شکل (۱۸) ساختار EAM در شبکه RIDNet. تفاوت رنگ لایه‌های کانولوشن به معنی تفاوت نرخ توسعه لایه کانولوشن است. لایه کانولوشن کوچک‌تر، کانولوشن 1×1 است [۹۷]



شکل (۱۹) معماری شبکه عصبی ADNet به تفکیک بلوک‌های تشکیل دهنده [۵۵]

است و می‌تواند به صورت مکمل یکدیگر عمل کرده و موجب بهبود عملکرد حذف نویز شبکه عصبی شود. ویژگی‌های استخراج شده توسط دو شاخه با یکدیگر ادغام شده و در اختیار لایه بازسازی تصویر قرار می‌گیرد. در آموزش این شبکه از رویکرد یادگیری رسوبی استفاده شده و این شبکه نویز تصویر را پیش‌بینی می‌کند. برای حذف پدیده شبکه بندی^۵ در HDRAB از کانولوشن توسعه یافته ترکیبی [۱۱۰] استفاده شده است. چون لایه‌های کانولوشن فقط می‌توانند اطلاعات محلی را استخراج کنند و نمی‌توانند از اطلاعات غیر محلی بهره‌برداری کنند و این موضوع ممکن است منجر به کاهش عملکرد حذف نویز شود، در RAB و HDRAB از مکانیسم توجه برای ثبت و یادگیری اطلاعات سراسری استفاده شده است. ساختار بلوک RAB در شکل (۲۱) نشان داده شده است. RAB شامل چند لایه کانولوشن متصل فعال‌ساز ReLU است که با اتصالات رسوبی به یکدیگر متصل شده‌اند. همچنین این بلوک شامل یک ماژول توجه فضایی^۶ (SAM) است که با استفاده از عملگر ادغام میانگین سراسری (GAP)، عملگر ادغام بیشینه سراسری^۷ (GMP) به همراه لایه کانولوشن و توابع فعال‌ساز ReLU و Sigmoid پیاده‌سازی شده است. در این بلوک نسبت به بلوک‌های رسوبی استاندارد، از اتصالات رسوبی بیشتری استفاده شده است تا عملکرد حذف نویز بلوک را بهبود بخشد.

DRANet شامل دو شاخه موازی با معماری متفاوت است که ویژگی‌های مکمل را استخراج کرده و موجب بهبود عملکرد مدل می‌شود. زیرشبکه اول شامل ۵ بلوک توجه رسوبی^۱ (RABs) با چندین اتصالات پرشی است و زیرشبکه دوم شامل ۵ بلوک توجه رسوبی هیبریدی توسعه یافته^۲ (HDRAB) است که در آن نیز از اتصالات پرش استفاده شده است. ساختار این شبکه عصبی را در شکل (۲۰) نمایش داده شده است. اتصالات پرشی ویژگی‌های سطح پایین، که شامل جزئیات تصویر ورودی هستند را به لایه‌های عمیق‌تر شبکه منتقل می‌کند. با این معماری، لایه‌های پایانی شبکه به ترکیبی از ویژگی‌های با انتزاع بالا و پایین دسترسی دارند و قادر هستند جزئیات تصویر ورودی را به خوبی بازیابی کنند. با استفاده از این دو بلوک، ویژگی‌های محلی با ارزش از طریق اتصالات پرشی چندگانه بین لایه‌های کانولوشن مختلف استخراج شده و ویژگی‌های بی‌اهمیت توسط ماژول توجه رسوبی حذف می‌شود. در این شبکه از عملیات نمونه‌برداری کاهش^۳ و کانولوشن توسعه یافته برای افزایش اندازه میدان پذیرش^۴ استفاده شده است، تا شبکه عصبی اطلاعات زمینه بیشتری از تصاویر استخراج کند. هر یک از شاخه‌ها معماری متفاوتی دارند و هر یک از مکانیزم توجه متفاوتی بهره می‌برند، بنابراین ویژگی‌های استخراج شده توسط هر یک از شاخه‌ها با شاخه دیگر متفاوت

¹ Residual Attention Block

² Hybrid Dilated Residual Attention Block

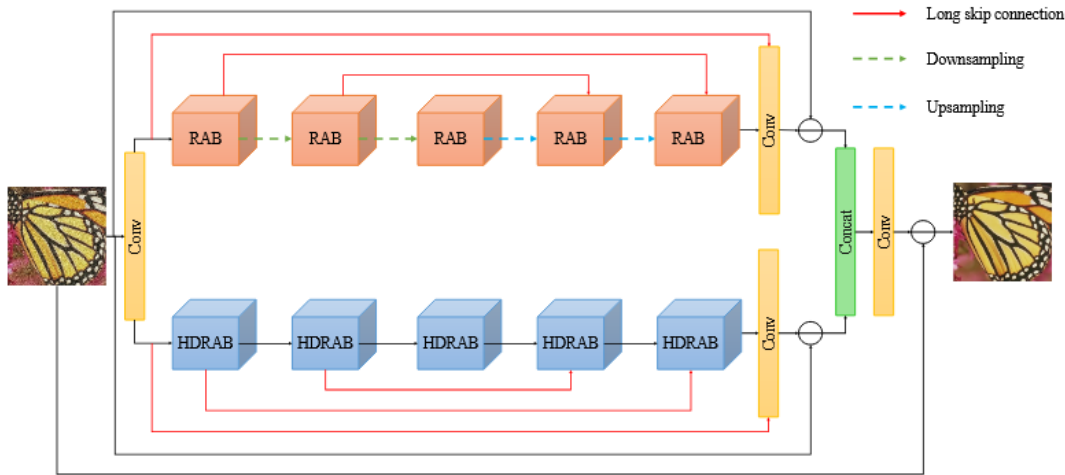
³ Downsampling

⁴ Receptive Field

⁵ Gridding Phenomenon

⁶ Spatial Attention Module

⁷ Global Max Pooling



شکل (۲۰) معماری شبکه عصبی DRANet [۹۲]

پراکنده، در SRB از اتصالات پرشی (با عملگر الحاق) استفاده شده تا با تسهیل جریان اطلاعات و خطا، ویژگی‌های استخراج شده را بهبود داده و از حذف اطلاعات زمینه جلوگیری کند. NEB برای استخراج نویز تصویر ورودی است. NEB از هفت لایه کانولوشن تشکیل شده است تا میدان پذیرش شبکه را افزایش دهد و به صورت موثرتر اطلاعات نویز را استخراج کند. برای بهبود جریان اطلاعات و ویژگی‌های استخراج شده، در NEB از دو اتصال رسوبی نیز استفاده شده است. به جای استفاده از مکانیزم توجه ترکیبی (توجه فضایی و توجه مبتنی بر کانال) ماژول توجه MADRNet با استفاده از یک شبکه دو شاخه به نام DRAB ایجاد شده است. DRAB از ترکیب کانولوشن معمولی و توسعه یافته تشکیل شده است و دارای دو زیر شاخه با طول متفاوت است و این دو شاخه با ساختاری مشابه ساختار رسوبی با یکدیگر ترکیب شده‌اند.

ماژول توجه در انتهای شبکه قرار دارد و در زمان بازسازی تصاویر با تمرکز بر روی ویژگی‌های مهم، به استخراج موثر اطلاعات نویز پنهان کمک می‌کند. خروجی لایه‌های NEB و SRB بر روی DRAB اعمال می‌شود تا به طور موثر از دست رفتن اطلاعات ویژگی‌های کم عمق جلوگیری کند و اثر حذف نویز را بهبود بخشد.

ژانگ و همکاران [۱۱۲] با معرفی ماژول توجه مبتنی بر هسته^۷ (KBA)، یک شبکه عصبی چند شاخه برای بازیابی تصاویر ارائه کردند که در آن، برای تولید هر پیکسل تصویر خروجی، مجموعه‌ای از فیلترهای قابل یادگیری مشارکت دارند. ماژول KBA وظیفه پیش‌بینی ضرایب ترکیب فیلترها را بر عهده دارد. ضرایب از این فیلترها بر روی تصویر ورودی (یا نقشه‌های ویژگی) اعمال می‌شود. با استفاده از KBA برای هر پیکسل از تصویر، یک فیلتر خاص متناسب با محتوای تصویر ایجاد می‌شود.

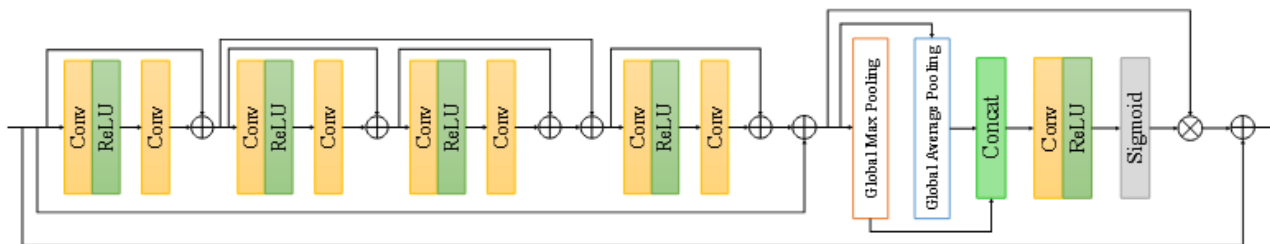
ساختار بلوک HDRAB در شکل (۲۲) نشان داده شده است. HDRAB از چندین لایه کانولوشن توسعه یافته ترکیبی با گام‌های متفاوت (بین ۱ تا ۴) تشکیل شده است که با اتصالات رسوبی متعدد به یکدیگر متصل شده‌اند. HDRAB شامل یک ماژول توجه مبتنی بر کانال^۱ (CAM) است که برای استخراج روابط بین کانالی موجود در ویژگی‌های کانولوشن استفاده شده و با استفاده از عملگر ادغام میانگین سراسری (GAP)، کانولوشن و توابع فعال‌ساز ReLU و Sigmoid پیاده‌سازی شده است.

دو و همکاران [۱۱۱] شبکه عصبی حذف نویز رسوبی با توجه چند مقیاسی^۲ به نام MADRNet ارائه کردند که از اتصالات پرشی، کانولوشن‌های توسعه یافته، مکانیزم پراکنده، مکانیزم توجه و رویکرد یادگیری تصویر رسوبی (نویز) استفاده می‌کند. MADRNet از ترکیب بلوک انتقال بازه‌ای متراکم^۳ (DTB)، بلوک رسوبی پراکنده^۴ (SRB)، بلوک بازسازی توجه رسوبی توسعه یافته^۵ (DRACB) و بلوک استخراج نویز^۶ (NEB) تشکیل شده است. معماری MADRNet در شکل (۲۳) نشان داده شده است. برای بهره‌مندی از مزایای معماری متراکم، DTB از نوع خاصی از معماری متراکم به نام بلوک متراکم بازه‌ای استفاده می‌کند که در آن یک لایه کانولوشن و یک لایه کانولوشن توسعه یافته، به صورت متوالی در کنار یکدیگر قرار دارند و ورودی هر لایه از ادغام خروجی دو لایه قبلی به دست می‌آید. این ساختار میدان پذیرش را افزایش داده (کانولوشن توسعه یافته) و در عین حال تعداد پارامترها را کاهش داده و موجب استخراج ویژگی‌های بهتر و کارآمدتر شده است. مکانیزم پراکنده در SRB با ترکیب لایه‌های کانولوشن معمولی و لایه‌های کانولوشن توسعه یافته پیاده‌سازی شده است. برای بهبود ساختار

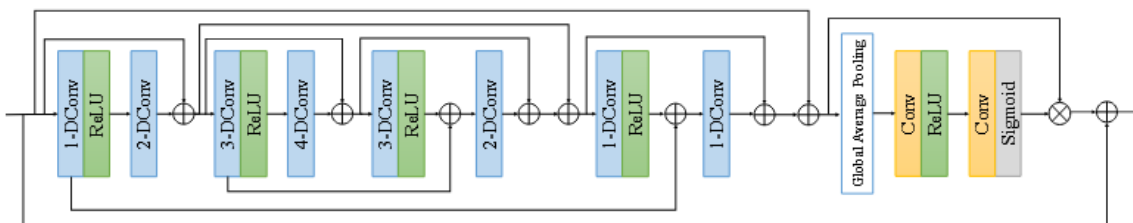
1 Channel Attention Module
 2 Multi-scale Attention Dilated Residual Network
 3 Dense Interval Transmission Block
 4 Sparse Residual Block
 5 Dilated Residual Attention Construction Block
 6 Noise Extraction Block

⁷ Kernel Basis Attention

در شکل (۲۴) نحوه محاسبه خروجی ماژول توجه مبتنی بر هسته نمایش داده شده است.



شکل (۲۱) معماری بلوک توجه رسوبی (RAB) در شبکه عصبی DRANet [۹۲]



شکل (۲۲) معماری بلوک توجه رسوبی هیبریدی توسعه یافته (HDRAB) در شبکه عصبی DRANet [۹۲]

ترتیب با ادغام حداکثری و درون‌یابی خطی انجام شده است. در زیر شاخه دیگر از کانولوشن توسعه یافته ترکیبی [۱۱۰] برای افزایش میدان پذیرش و کاهش پیچیدگی و همچنین حذف پدیده شبکه‌ای^۲ استفاده شده است. در این شبکه از مکانیزم توجه مبتنی بر کانال در کنار مکانیزم توجه فضایی استفاده شده است. معماری ماژول توجه DCANet در شکل (۲۵) نشان داده شده است. ماژول توجه مبتنی بر کانال با استفاده از GAP و دو لایه کانولوشن پیاده‌سازی شده است. ماژول توجه فضایی با استفاده از ترکیب GAP و GMP پیاده‌سازی شده و خروجی این دو عملگر بعد از عبور از یک لایه کانولوشن برای وزن‌دهی به نقشه ویژگی استفاده شده‌اند. ماژول توجه در این شبکه در ابتدای شبکه و پیش از زیر شبکه حذف نویز قرار دارد و به صورت یک ماژول رسوبی در شبکه قرار گرفته است.

۴- مقایسه شبکه‌های عصبی حذف نویز

شبکه‌های عصبی عمیق بر پایه مدل‌های آماری غیرخطی بسیار پیچیده و با پارامترهای بسیار زیاد هستند و این موضوع شفافیت آن‌ها را با چالش مواجه کرده است. به عبارت دیگر شبکه‌های عصبی ساختار جعبه سیاه دارند و توضیح نحوه عملکرد آن‌ها و ارائه دلیل یا توضیح مناسب برای تصمیم‌ها و عملکرد این شبکه‌ها به طور کامل و دقیق برای انسان ممکن نیست [۱۱۳].

با وجود پژوهش‌هایی در حوزه شبکه‌های عصبی قابل توضیح^۳، بررسی معماری شبکه‌های عصبی حذف نویز اشاره شده در این پژوهش از نظر ساختارهای مشترک مورد استفاده، می‌تواند راهگشای پژوهش‌های آینده حذف نویز تصویر باشد. به همین منظور در این بخش به صورت خلاصه به بررسی ساختار کلی

F ماتریس ضرایب فیلترهای قابل یادگیری است که در فرآیند آموزش، مقادیر آن به نحوی بهینه شده است که به صورت تطبیق‌پذیر، بر اساس تصویر ورودی، ترکیب مناسبی از فیلترهای قابل یادگیری را تولید می‌کند. W مجموعه‌ای از فیلترهای قابل یادگیری پایه که بین تمام تصاویر و تمام موقعیت‌های مکانی تصاویر مختلف مشترک هستند را نمایش می‌دهد. این فیلترها به نحوی آموزش داده شدند تا بتوانند الگوهای مشترک و رایج تصاویر را استخراج کنند. برای هر نقطه از تصویر یا نقشه ویژگی ورودی، با مختصات (i,j) ، یک فیلتر تطبیقی $M(i,j)$ با ضرب، F در W تولید می‌شود. X نقشه ویژگی ورودی را نمایش داده و X_e نقشه ویژگی بهبود یافته است که با اعمال کانولوشن 1×1 بر روی X محاسبه می‌شود. اعمال این کانولوشن برای این است که نقشه‌های ویژگی متناسب با فیلتر کانولوشن تطبیقی M تولید شود. در این روش از ترکیب توجه فضایی و مبتنی بر کانال استفاده شده است. از U-Net [۶۲] به عنوان معماری پایه این شبکه عصبی استفاده شده است. برای ترکیب ویژگی‌های متنوع به منظور بهبود فرآیند بازیابی تصویر، یک بلوک چند شاخه به نام ترکیب چند محوری ویژگی‌ها^۱ (MFF) برای مدیریت اطلاعات فضایی و مبتنی بر کانال ارائه شده است. این روش توانایی بالایی در مسائل مختلف بازیابی تصویر از خود نشان داده ولی در عین حال پیچیدگی ساختاری بالایی دارد.

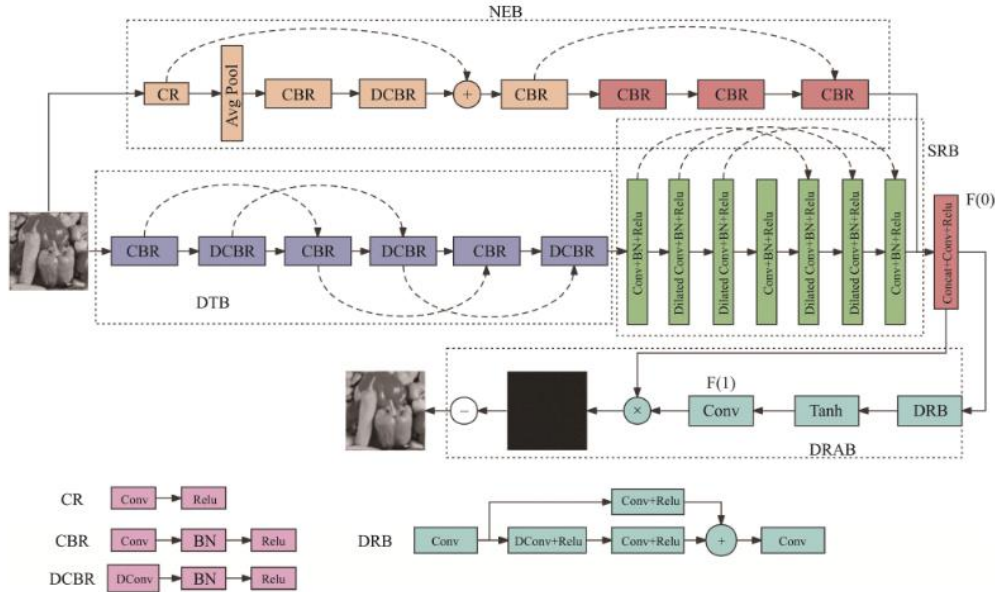
وو و همکاران، شبکه عصبی عمیق DCANet [۹۵] را برای حذف نویزهای واقعی ارائه کردند. این شبکه دارای یک زیر شبکه عصبی ۷ لایه برای تخمین نویز است و یک زیر شبکه دوشاخه، فرآیند بازیابی نویز، از تصویر نویزی را انجام می‌دهد. هر یک از شاخه‌های زیر شبکه بازیابی نویز، دارای ساختاری متفاوت هستند و اطلاعات مکملی استخراج می‌کنند. یکی از زیر شاخه‌ها دارای معماری U شکل است و فرآیند نمونه‌برداری کاهشی و افزایشی به

² Gridding Phenomenon

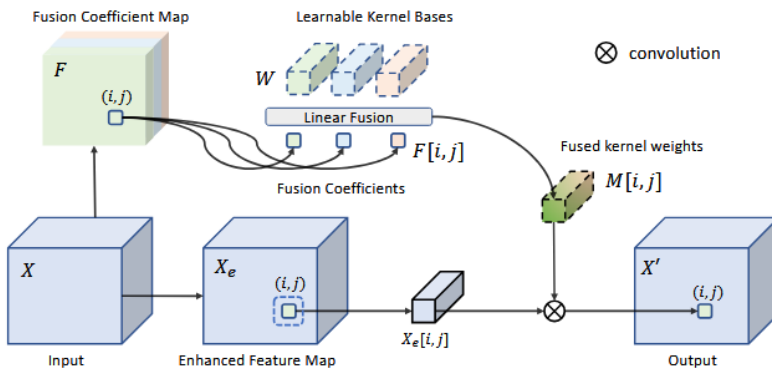
³ Explainable Neural Networks

¹ Multi-axis Features Fusion

شبکه‌های حذف نویز و هم‌چنین چالش‌های طراحی و پیاده‌سازی این شبکه‌ها می‌پردازیم.



شکل (۲۳) ساختار شبکه عصبی حذف نویز MADRNet [۱۱۱].



شکل (۲۴) ساختار ماژول توجه مبتنی بر هسته در شبکه عصبی KBNet [۱۱۲].

بتواند تصویر بازسازی شده را تولید کند. شبکه‌های چند منظوره بازبازی تصویر معمولاً برای انجام وظایف مختلف بازبازی، از بلوک‌های بازسازی مختلف استفاده می‌کنند. (برای وظیفه حذف نویز از یک بلوک بازسازی و برای هر یک از وظایف وضوح فوق العاده تصویر و حذف مصنوعات فشرده‌سازی JPG، از بلوک‌های بازسازی متفاوت)

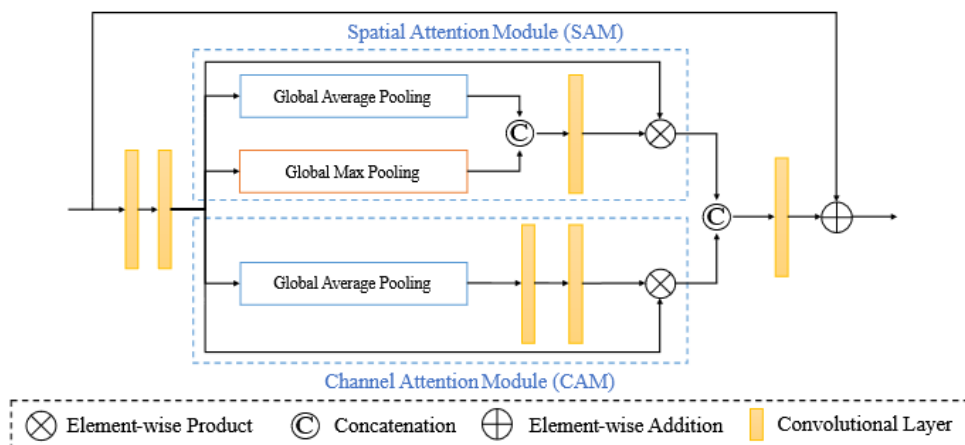
یکی از رویکردهای رایج در شبکه عصبی حذف نویز آموزش رسوبی است. در این رویکرد به جای اینکه شبکه عصبی الگوهای تصویر را بیاموزد، الگوهای نویز را یاد می‌گیرد. به عبارت دیگر شبکه عصبی می‌آموزد که در لایه‌های پنهان خود، الگوهای تصویر را حذف کند و نویز تصویر را به عنوان خروجی تولید نماید. این رویکرد در بسیاری از پژوهش‌ها [۵، ۵۵، ۶۱، ۷۹، ۸۷، ۹۲، ۹۵، ۹۷، ۱۱۱] مورد استفاده قرار گرفته و موجب بهبود یادگیری شبکه و افزایش کارایی حذف نویز شبکه عصبی شده است. پیاده‌سازی این رویکرد با یک اتصال کوتاه انجام می‌شود. این اتصال کوتاه در شکل (۲۶) به صورت خط ممتد مشخص شده است و تصویر ورودی را به انتهای شبکه عصبی منتقل می‌کند. با کم کردن مقدار

۴-۱- ساختار کلی شبکه‌های عصبی حذف نویز مدرن

ساختار کلی بسیاری از شبکه‌های عصبی حذف نویز مدرن [۵، ۵۵، ۹۲، ۹۵، ۹۷، ۱۱۱] از معماری شکل (۲۶) پیروی می‌کند. معماری تعدادی از شبکه‌های عصبی مدرن [۶۴-۶۶، ۱۱۲] نیز بر پایه معماری U-Net طراحی شده است که این معماری نیز به طور کلی ساختاری مشابه شکل (۲۶) دارد.

شبکه‌های عصبی حذف نویز از چند بلوک اصلی تشکیل شده‌اند. بلوک استخراج ویژگی کم عمق که در ابتدای شبکه عصبی قرار دارد و معمولاً از یک لایه کانولوشن و فعال‌ساز (معمولاً ReLU) تشکیل شده است. این بلوک تصاویر ورودی را (با عمق یک برای تصاویر خاکستری یا سه برای تصاویر رنگی) به فضای ویژگی‌های پنهان نگاشت می‌کند. معمولاً ویژگی‌های فضای پنهان عمق بیشتری دارد. یکی دیگر از بلوک‌های تشکیل دهنده شبکه‌های عصبی حذف نویز بلوک بازسازی تصویر است. این بلوک نیز یک بلوک کم عمق است که در انتهای شبکه عصبی قرار دارد و تنها شامل یک لایه کانولوشن است که فاقد تابع فعال‌ساز است تا

نویز پیش‌بینی شده از تصویر نویزدار، تصویر بازیابی شده تمیز تولید می‌شود.



شکل (۲۵) ماژول توجه شبکه عصبی حذف نویز DCANet. ماژول توجه مبتنی بر کانال (CAM) و ماژول توجه فضایی (SAM) با خط چین مشخص شده است [۹۵]

بازیابی جزئیات و بافت‌های پیچیده تصویر در بازیابی تصویر از اهمیت ویژه‌ای برخوردار است. این ویژگی‌ها ممکن است در لایه‌های عمیق شبکه عصبی از بین بروند. برای جلوگیری از این موضوع و انتقال جزئیات تصویر به لایه‌های پایانی، ویژگی‌های کم عمق شبکه عصبی (ویژگی‌های سطح پایین شامل لبه‌ها، گوشه‌ها و ...) با استفاده از یک اتصال کوتاه به ورودی لایه بازسازی تصویر منتقل شده و امکان استفاده از این جزئیات برای لایه بازسازی تصویر فراهم می‌شود. این اتصال کوتاه در شکل (۲۶) به صورت خط‌چین نمایش داده شده است و در پژوهش‌های مختلف [۷۹، ۸۲، ۸۴، ۸۵، ۸۷، ۹۲، ۹۵، ۹۷، ۱۰۶، ۱۰۷، ۱۱۴] مورد استفاده قرار گرفته است.

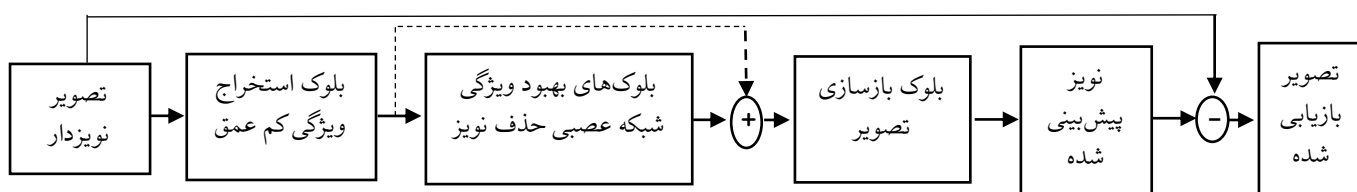
بلوک بهبود ویژگی در شبکه‌های عصبی مختلف حذف نویز ساختار متفاوتی دارد. این بخش از شبکه عصبی در شکل (۲۶) به صورت خط‌چین نمایش داده شده است و در پژوهش‌های مختلف [۷۹، ۸۲، ۸۴، ۸۵، ۸۷، ۹۲، ۹۵، ۹۷، ۱۰۶، ۱۰۷، ۱۱۴] مورد استفاده قرار گرفته است.

بلوک بهبود ویژگی در شبکه‌های عصبی مختلف حذف نویز ساختار متفاوتی دارد. این بخش از شبکه عصبی در شکل (۲۶) به صورت خط‌چین نمایش داده شده است و در پژوهش‌های مختلف [۷۹، ۸۲، ۸۴، ۸۵، ۸۷، ۹۲، ۹۵، ۹۷، ۱۰۶، ۱۰۷، ۱۱۴] مورد استفاده قرار گرفته است.

بلوک بهبود ویژگی در شبکه‌های عصبی مختلف حذف نویز ساختار متفاوتی دارد. این بخش از شبکه عصبی در شکل (۲۶) به صورت خط‌چین نمایش داده شده است و در پژوهش‌های مختلف [۷۹، ۸۲، ۸۴، ۸۵، ۸۷، ۹۲، ۹۵، ۹۷، ۱۰۶، ۱۰۷، ۱۱۴] مورد استفاده قرار گرفته است.

بلوک بهبود ویژگی در شبکه‌های عصبی مختلف حذف نویز ساختار متفاوتی دارد. این بخش از شبکه عصبی در شکل (۲۶) به صورت خط‌چین نمایش داده شده است و در پژوهش‌های مختلف [۷۹، ۸۲، ۸۴، ۸۵، ۸۷، ۹۲، ۹۵، ۹۷، ۱۰۶، ۱۰۷، ۱۱۴] مورد استفاده قرار گرفته است.

بلوک بهبود ویژگی در شبکه‌های عصبی مختلف حذف نویز ساختار متفاوتی دارد. این بخش از شبکه عصبی در شکل (۲۶) به صورت خط‌چین نمایش داده شده است و در پژوهش‌های مختلف [۷۹، ۸۲، ۸۴، ۸۵، ۸۷، ۹۲، ۹۵، ۹۷، ۱۰۶، ۱۰۷، ۱۱۴] مورد استفاده قرار گرفته است.



شکل (۲۶) ساختار کلی شبکه‌های عصبی حذف نویز مدرن

در معماری‌های مختلف شبکه عصبی امکان بهره‌برداری از مبدل Swin وجود دارد. یادگیری رسوبی نیز در هر چهار شبکه عصبی مورد استفاده قرار گرفته است. شبکه‌های عصبی NLRN [۶۹] و MWCNN [۴۳] با ایجاد تغییر در عملکرد اصلی شبکه عصبی، موجب بهبود عملکرد شبکه عصبی شده‌اند. در NLRN، امکان بهره‌برداری از ویژگی‌های غیر محلی برای شبکه عصبی فراهم شده است. شبکه‌های عصبی کانولوشن ذاتاً از ویژگی‌های محلی بهره‌برداری می‌کنند. ترکیب ویژگی‌های محلی و غیر محلی در NLRN موجب بهبود عملکرد حذف نویز شده است. در MWCNN نیز، جایگزینی تبدیل موجک و معکوس آن با عملگر نمونه‌برداری کاهشی و افزایشی موجب بهبود حفظ جزئیات تصویر شده است.

۴-۲- طراحی و پیاده‌سازی شبکه‌های عصبی حذف نویز

شبکه‌های عصبی کانولوشن در سال‌های اخیر موفقیت‌های چشمگیری را در زمینه حذف نویز تصاویر ارائه کرده‌اند. مدل‌های مبتنی بر CNN بسیار عمیق، ظرفیت و انعطاف پذیری بسیار بالایی برای بهره‌برداری از ویژگی‌های تصویر ارائه کردند. با این حال پیاده‌سازی و آموزش CNN‌های عمیق همواره با مشکلاتی همراه است. آموزش یک CNN بسیار عمیق برای وظایف حذف نویز بسیار دشوار است و اکثر CNN‌های عمیق از اشباع عملکرد (گرادیان محوشونده یا انفجاری)^۲ رنج می‌برند [۱۱۵, ۷۷, ۵۰]. در شبکه‌های عصبی عمیق توزیع ورودی‌های هر لایه در طول آموزش تغییر می‌کند و مقابله با این موضوع نیازمند نرخ یادگیری پایین‌تر و مقداردهی اولیه دقیق پارامترها است و این موضوع آموزش را کند می‌کند و آموزش مدل‌هایی با توابع غیرخطی اشباع شده را به طرز چشمگیری دشوار می‌سازد [۵۰]. (به این پدیده، تغییر متغیر داخلی^۳ گفته می‌شود)

در سال‌های اخیر پیشرفت‌های قابل توجهی در روش‌های آموزش شبکه‌های عصبی به دست آمده است که آموزش شبکه‌های عصبی عمیق را بهبود داده است. تعدادی از این موارد عبارتند از: واحد یکسوساز خطی^۴ (ReLU) [۴۹]، نرمال‌سازی دسته‌ای^۵ [۵۰] و یادگیری رسوبی^۶ [۵۱]. این روش‌ها می‌توانند باعث سرعت بخشیدن به روند آموزش و همچنین بهبود عملکرد حذف نویز در شبکه‌های عصبی عمیق حذف نویز شوند. استفاده از تابع فعال‌ساز ReLU به عنوان جایگزین توابع فعال‌سازی که در فرآیند آموزش با مشکل اشباع مواجه می‌شوند (مثل تابع سیگموئید) می‌تواند فرآیند آموزش شبکه‌های عصبی عمیق را تسهیل کند. نرمال‌سازی دسته‌ای برای مقابله با مشکل تغییر متغیر داخلی ارائه

معیارهای ارزیابی روش‌های حذف نویز و مقایسه عملکرد شبکه‌های عصبی مختلف در حذف نویز تصاویر در بخش ۴-۴ ارائه شده است. همانطور که در آن بخش اشاره شده، همانطور که در آن بخش اشاره شده، سه شبکه عصبی KBNNet [۸۸]، SCUNet [۴۴] و SwinIR [۸۲] در بازیابی تصاویر رنگی و خاکستری تخریب شده با سطوح نویز مختلف بهترین نتایج را ارائه داده‌اند. بعد از این سه شبکه عصبی، شبکه‌های عصبی NLRN [۶۹] و MWCNN [۴۳] در بازیابی تصاویر خاکستری نسبت به سایر شبکه‌های عصبی، عملکرد بهتری ارائه کردند و شبکه عصبی DRANet [۹۲] نیز در حذف نویز تصاویر رنگی با سطوح نویز مختلف نتایج بهتری نسبت به سایر شبکه‌های عصبی ارائه داده است. همانطور که در جدول (۱) آورده شده است، چهار شبکه عصبی KBNNet [۸۸]، SCUNet [۴۴]، SwinIR [۸۲] و DRANet [۹۲] از نظر ساختاری، از مکانیزم توجه مبتنی بر مبدل‌ها یا مکانیزم خود توجه چند سر^۱ (MSA) است که از چندین مکانیزم خود توجه موازی بهره می‌برد. مکانیزم توجه KBNNet و DRANet نیز ترکیبی از مکانیزم توجه کانال و مکانیزم توجه فضایی است. (مکانیزم توجه فضایی این دو شبکه با هم متفاوت است) می‌توان نتیجه گرفت که استفاده از چندین مکانیزم توجه مختلف در کنار یکدیگر در وظایف بازیابی تصویر می‌تواند موجب بهبود ویژگی‌های استخراج شده و در نهایت بهبود عملکرد شبکه عصبی شود. برخلاف برخی شبکه‌های عصبی که از مکانیزم توجه، تنها یکبار و در انتهای شبکه عصبی بهره می‌برند، این چهار شبکه عصبی، در ویژگی‌های سطوح مختلف انتزاع شبکه عصبی (ویژگی‌های لایه‌های مختلف) از مکانیزم توجه استفاده کرده و ویژگی‌های سطوح مختلف را با استفاده از توجه بهبود داده‌اند. استفاده از چنین ساختاری نیز ممکن است موجب بهبود عملکرد شبکه‌های عصبی حذف نویز شود. علاوه بر این، محاسبه وزن‌ها در مکانیزم توجه مورد استفاده در سه شبکه عصبی KBNNet، SCUNet و SwinIR به صورت جهانی نیست و وزن‌های توجه در یک همسایگی از هر پیکسل محاسبه می‌شود، بنابراین محدود کردن محدوده مکانیزم توجه به یک همسایگی از پیکسل‌ها ممکن است موجب بهبود عملکرد حذف نویز شود. با توجه به اینکه SCUNet و SwinIR از مبدل Swin استفاده می‌کنند، استفاده از مبدل Swin در ساختار شبکه عصبی موجب بهبود قابل توجه عملکرد شبکه عصبی حذف نویز می‌شود. با توجه به اینکه معماری کلی این دو شبکه عصبی متفاوت است و یکی از آن‌ها از ساختار شبکه U-Net پیروی می‌کند و دیگری از ساختاری عمیق متشکل از بلوک‌های رسوبی متوالی تشکیل شده است، می‌توان نتیجه گرفت که بهبود عملکرد حاصل از ترکیب مبدل Swin با شبکه عصبی محدود به ساختاری خاصی نیست و

² Vanishing or Exploding Gradients

³ Internal Covariate Shift

⁴ Rectifier Linear Unit

⁵ Batch Normalization

⁶ Residual Learning

¹ Multi Head Self Attention

۸۴، ۹۲، ۹۵، ۱۱۲، ۱۱۴]. برای این منظور تصاویر مجموعه‌های داده تصویری معروف در ابعاد مشخص و به صورت وصله‌های همپوشان برش داده شده و سپس انواع تغییرات مانند چرخش، تغییر مقیاس و آینه‌ای کردن^۱ بر روی آن‌ها اعمال می‌شود تا نمونه‌های داده متنوعی تولید شود. با این کار شبکه عصبی حذف نویز با الگوهای متنوع تری آموزش داده شده و عملکرد شبکه عصبی حذف نویز بهبود می‌یابد.

علاوه بر موارد بیان شده در این بخش که مربوط به چالش‌های طراحی و آموزش شبکه‌های عصبی حذف نویز است، یکی از مهمترین چالش‌های آموزش شبکه‌های عصبی عمیق حذف نویز، هزینه پردازشی بالای آموزش این شبکه‌ها است. آموزش این شبکه‌ها با تنها استفاده از GPUهای قدرتمند امکان پذیر است و زمان بسیار زیادی به طول می‌انجامد. با توجه به اینکه بخشی از فرآیند ارزیابی شبکه‌های عصبی بررسی اثربخشی بخش‌ها و ماژول‌های مختلف معماری پیشنهادی به صورت جزء به جزء است و کارایی شبکه عصبی در صورت وجود یا عدم وجود یک ماژول (برای مثال یک ماژول توجه، یا یک اتصال رسوبی و ...) مورد بررسی قرار می‌گیرد، هزینه پردازشی بالای آموزش شبکه عصبی بیش از پیش اهمیت پیدا کرده و دستیابی به یک مدل بهینه را با چالش مواجه می‌کند. برای مثال آموزش مدل حذف نویز SCUNet [۶۴] بر روی پردازنده گرافیکی NVIDIA RTX 2080 Ti به مدت سه روز، آموزش مدل حذف نویز DRANet [۹۲] بر روی پردازنده گرافیکی NVIDIA GeForce RTX 3080Ti به مدت ۱۹۰ ساعت و آموزش مدل حذف نویز DCANet [۹۵] بر روی پردازنده گرافیکی NVIDIA GeForce RTX 3080Ti به مدت ۴۸ ساعت به طول می‌انجامد. با توجه به این موارد، دسترسی به GPUهای گران قیمت یکی از ضرورت‌های آموزش شبکه‌های عصبی حذف نویز است.

۴-۳- آموزش و ارزیابی شبکه‌های عصبی حذف نویز

برای آموزش مدل‌های حذف نویز از مجموعه‌های تصویر طبیعی بزرگ و برچسب‌گذاری نشده مختلف مانند نسخه‌های مختلف Berkeley Segmentation Dataset (BSDS400) و BSD500 [۱۱۶]، زیر مجموعه‌های ImageNet [۱۱۷] و Waterloo Exploration Database [۱۱۸] استفاده شده است. برای آموزش مدل‌های حذف نویز گوسی تصویر، تصاویر مجموعه‌های داده به صورت دستی با نویز گوسی سفید جمع شوند (AWGN) تخریب شده و تصویر تخریب شده و تصویر تمیز به مدل اعمال می‌شود. همانطور که در بخش ۴-۲ اشاره شد استفاده از تکنیک‌های مختلف افزایش داده نیز در آموزش

شده است و به طور چشمگیری منجر به افزایش سرعت آموزش شبکه عصبی، بهبود عملکرد حذف نویز و همچنین کاهش حساسیت فرآیند آموزش به مقادیر اولیه شده است [۵۰، ۱۱۵]. علاوه بر این استفاده از توالی لایه کانولوشن و نرمال‌سازی دسته‌ای در معماری شبکه‌های عصبی حذف نویز عملکرد بسیار خوبی در حذف نویز تصویر ارائه کرده است [۵]. اگرچه افزایش عمق شبکه می‌تواند عملکرد حذف نویز تصویر را بهبود بخشد، اما شبکه عمیق‌تر ممکن است منجر به مشکلات گرادینان محو شونده یا انفجار شود. یادگیری رسوبی [۵۱] ابزار خوبی برای حل این مشکل است. در این روش ورودی یک بلوک (شامل چندین لایه کانولوشن و فعال‌ساز) و خروجی آن با یکدیگر جمع شده و به عنوان ورودی لایه بعدی اعمال می‌شود تا افزایش کارایی را تضمین کند. گسترش عرض شبکه عصبی نیز راهکاری مفید و کارآمد به عنوان جایگزینی برای افزایش عمق شبکه و بهبود عملکرد حذف نویز است که در سال‌های اخیر و در مسائل مختلف، از جمله حذف نویز تصاویر [۵۹، ۷۵-۸۸] مورد استفاده قرار گرفته است. این موضوع در بخش ۳-۴ مورد بررسی قرار گرفت. افزایش عرض شبکه ممکن است پارامترهای شبکه را به شدت افزایش داده و پیچیدگی مدل حذف نویز را افزایش دهد. برای پاسخ به این مشکلات، کانولوشن‌های توسعه یافته ارائه شدند [۵۵]. کانولوشن توسعه یافته موجب افزایش میدان پذیرش شبکه عصبی، بدون افزایش عمق شبکه می‌شود و توانایی شبکه را برای ثبت اطلاعات بیشتر تصویر گسترش می‌دهد، ولی ممکن است موجب پدیده شبکه بندی^۱ شود. کانولوشن توسعه یافته ترکیبی [۱۱۰] برای رفع این مشکل ارائه شده است. نحوه استفاده از ترکیب پیشرفت‌های حاصل شده در حوزه طراحی و آموزش شبکه‌های عصبی و نحوه قراردادی ماژول‌ها یا لایه‌های مختلف (نرمال‌سازی دسته‌ای، فعال‌ساز ReLU و کانولوشن توسعه یافته)، مکان استفاده از اتصالات رسوبی، چگونگی افزایش عرض شبکه و معماری هر یک از شاخه‌های شبکه عصبی و نحوه ترکیب کانولوشن توسعه یافته و معمولی در معماری شبکه عصبی از چالش‌های طراحی شبکه‌های عصبی حذف نویز است که بخش ۴-۱ برای راهنمایی در خصوص این چالش‌ها ارائه شده است.

آموزش شبکه‌های عصبی عمیق به خصوص زمانی که نمونه‌های آموزشی محدودی در دسترس باشد، بسیار مشکل است. یک راه حل این مشکل استفاده از یک شبکه از پیش آموزش دیده و سپس تنظیم دقیق^۲ آن بر روی مجموعه داده هدف است. این روش در مسائل طبقه‌بندی تصویر کاربرد دارد ولی در روش‌های حذف نویز تصویر کاربرد چندانی ندارد. راه حل دیگر استفاده از افزایش داده^۳ گسترده و موثر است. این روش در آموزش شبکه‌های عصبی حذف نویز کاربرد بسیاری دارد [۵، ۲۰، ۵۶، ۵۹، ۶۴، ۷۷، ۷۸،

¹ Gridding Phenomenon

² Fine-Tuning

³ Data Augmentation

مبتنی بر مدل معروف و شبکه‌های عصبی مختلف حذف نویز ارائه شده است. پنج دقت برتر برای هر سطح نویز با زیر خط مشخص شده است. برای شبکه‌های عصبی که هم نسخه حذف نویز کور و هم غیر کور دارند، دقت حذف نویز کور در جدول درج شده است که کمی کمتر از دقت حذف نویز غیر کور است. همانطور که در جدول‌های (۲) تا (۴) مشخص است مشخص است استفاده از پیشرفت‌های حاصل در شبکه عصبی، به خصوص استفاده از شبکه‌های عصبی چند شاخه (که ویژگی‌های مختلفی را از تصویر ورودی استخراج می‌کنند)، مکانیزم توجه و بهره‌برداری از مشابهت‌های غیر محلی و استفاده از کانولوشن توسعه یافته تاثیر بسیاری در بهبود عملکرد شبکه عصبی در حذف نویز تصویر داشته است و شبکه‌های عصبی موفق در حذف نویز تصویر از یک یا ترکیبی از این موارد بهره برده‌اند. با توجه به ساختار جعبه سیاه شبکه عصبی بررسی نظری تاثیر هر یک از این عوامل بر بهبود عملکرد شبکه عصبی ممکن نیست ولی در هریک از این پژوهش‌ها با انجام آزمایشات عملی (مطالعات فرسایشی^۴)، تاثیر هر یک از این عوامل بر بهبود عملکرد شبکه عصبی مورد بررسی قرار گرفته است.

در یک مقایسه کلی بین روش‌های مبتنی بر مدل موفق با روش‌های مبتنی بر شبکه عصبی حذف نویز، بر اساس مقادیر جدول (۲)، در تمامی سطوح نویز همه روش‌های حذف نویز مبتنی بر شبکه عصبی به جز MLP و TNRD و DnCNN (تنها در نویز با $\sigma=75$) بر روی پایگاه داده SET12، بر روش‌های موفق مبتنی بر مدل (BM3D و WNNM) برتری دارند. بر اساس مقادیر جدول (۳)، تمام روش‌های مبتنی بر شبکه عصبی به جز BUIFD، در تمامی سطوح نویز بر روی تصاویر پایگاه داده BSD68، بر روش‌های موفق مبتنی بر مدل (BM3D و WNNM) برتری دارند. در حذف نویز تصاویر رنگی نیز، بر اساس جدول (۴) در تمامی سطوح نویز بر روی پایگاه داده CBSD68، تمام روش‌های مبتنی بر شبکه عصبی به جز BUIFD و DnCNN (تنها در نویز با $\sigma=75$) بر روش معیار و موفق مبتنی بر مدل CBM3D برتری دارد. این موضوع نشان دهنده ظرفیت بالای شبکه‌های عصبی در استخراج و بهره برداری از ویژگی‌های تصویر و برتری شبکه‌های عصبی حذف نویز، نسبت به روش‌های مبتنی بر مدل حذف نویز است.

سه شبکه عصبی حذف نویز KBNNet [۸۸]، SCUNet [۴۴] و SwinIR [۸۲] در بازیابی تصاویر رنگی و خاکستری تخریب شده با سطوح نویز مختلف، در هر سه مجموعه داده SET12، BSD68 و CBSD68 بهترین نتایج را نسبت به سایر روش‌ها ارائه داده‌اند. پس از سه شبکه عصبی فوق، شبکه عصبی NLRN [۶۹] و سپس شبکه عصبی MWCNN [۴۳] نیز در بازیابی تصاویر خاکستری با سطوح مختلف نویز، نسبت به سایر

مدل‌های حذف نویز کاربرد دارد [۵]، ۲۰، ۵۶، ۵۹، ۶۴، ۷۷، ۷۸، ۸۴، ۹۲، ۹۵، ۱۱۲، ۱۱۴].

برای ارزیابی روش پیشنهادی با AWGN مصنوعی، از حذف نویز تصویر بر روی تصاویر خاکستری، از مجموعه‌های داده BSD68 [۱۱۶] و Set12 [۵] استفاده شده است. مجموعه داده BSD68 شامل ۶۸ تصویر متفاوت از BSD300 است که دارای موضوعات متفاوت هستند. مجموعه داده Set12 مجموعه‌ای از تصاویر پرکاربرد در زمینه ارزیابی روش‌های حذف نویز است که شامل ۱۲ تصویر بوده و هر تصویر ابعاد 256×256 دارد. تصاویر این مجموعه‌های داده با AWGN با انحراف معیار استاندارد ۱۵، ۲۵، ۳۵، ۵۰ و ۷۵ تخریب شده و مدل حذف نویز با اعمال این برای ارزیابی حذف نویز تصویر رنگی، از مجموعه‌های داده، CBSD68 [۱۱۹] Kodak24، استفاده شده است. مجموعه داده CBSD68 نسخه رنگی مجموعه داده BSD68 است.

۴-۴- معیارهای ارزیابی روش‌های حذف نویز

برای ارزیابی روش‌های حذف نویز از معیارهای مختلفی استفاده می‌شود. اولین معیار، بررسی بصری تصاویر بازیابی شده است که نوعی معیار کیفی محسوب می‌شود. مهمترین معیارهای کمی ارزیابی روش‌های حذف نویز PSNR و SSIM است. PSNR یک معیار برای بیان حداکثر توان ممکن یک سیگنال نسبت به قدرت نویز مخرب آن است و از رابطه (۹) محاسبه می‌شود. PSNR ترکیب ساختاری تصویر را که به خوبی توسط بینایی انسان درک می‌شود، در نظر نمی‌گیرد.

$$PSNR = 10 * \log_{10} \frac{255^2}{MSE}$$

$$MSE = \frac{1}{MN} \sum_{m=0}^M \sum_{n=0}^N (Resetored[m, n] - clean[m, n])^2$$

(۹)
از نظر بینایی انسان، بهتر از PSNR است. SSIM یک معیار ارزیابی بر اساس ترکیب درخشندگی^۱، کنتراست^۲ و تغییرات در اطلاعات ساختاری^۳ است و از رابطه (۱۰) محاسبه می‌شود.

$$SSIM(x, y) = \frac{(2 * \mu_x \mu_y + c_1)(2 * \sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (10)$$

۴-۵- نتایج پیاده‌سازی روش‌های مختلف حذف نویز

در جدول (۲) و (۳) مقدار PSNR تصاویر بازیابی شده مجموعه داده SET12 و BSD68 با سطوح نویز مختلف با استفاده از روش‌های مبتنی بر مدل معروف و شبکه‌های عصبی حذف نویز ارائه شده است. در جدول (۴) نیز مقادیر PSNR تصاویر بازیابی شده پایگاه داده CBSD68 با استفاده از روش‌های

1 Luminance

2 Contrast

3 Structural Information

4 Ablation Study

تصاویر تخریب شده با بیشترین شدت نویز مرتب شده‌اند، مدل‌های سمت چپ نمودار، ضعیفترین عملکرد را (بر اساس معیار PSNR) در بازیابی تصاویر نویزی ارائه کرده‌اند و هر چه از سمت چپ نمودار به سمت راست حرکت کنیم، مدل‌های حذف نویز بر اساس معیار PSNR، عملکرد بهتری از خود نشان داده‌اند. مدل‌های سمت چپ نمودار شامل روش‌های مبتنی بر مدل و شبکه‌های عصبی DnCNN، IRCNN، MLP، BUIFD، ADNet، FFDNet، DudeNet و N3Net هستند که به جز DudeNet و BUIFD، سایر مدل‌ها، شبکه‌های عصبی تک شاخه هستند. همانطور که در بخش (۳-۴) گفته شد، برای اینکه شبکه‌های عصبی چند شاخه بتوانند ویژگی‌های بهتری را استخراج کنند، باید معماری شاخه‌های مختلف شبکه‌های عصبی با یکدیگر متفاوت باشد تا هر شاخه، ویژگی‌های متفاوتی را استخراج کند و ترکیب ویژگی‌های شاخه‌های مختلف موجب بهبود عملکرد شبکه عصبی چند شاخه شود. شبکه عصبی BUIFD در تمامی شاخه‌ها از یک معماری یکسان استفاده کرده و احتمالاً این موضوع باعث شده که این شبکه عصبی چند شاخه نتوانسته ویژگی‌های موثری از تصویر استخراج کند و عملکرد ضعیفی در حذف نویز تصویر، به خصوص در شدت نویز بالا از خود نشان دهد این شبکه عصبی در نویزهای شدید عملکردی ضعیفتر از روش‌های مبتنی بر مدل کرده است.

تنها شبکه عصبی چند شاخه، که معماری آن با اصول طراحی شبکه‌های چند شاخه سازگاری دارد و مانند سایر شبکه‌های عصبی چند شاخه، از معماری متفاوتی برای هر شاخه استفاده کرده ولی از نظر عملکرد در میان شبکه‌های عصبی تک شاخه نمودار قرار گرفته، شبکه عصبی DudeNet است. همانطور که در ابتدای این فصل گفته شد برای شبکه‌های عصبی که هم نسخه حذف نویز کور و هم غیر کور دارند، دقت حذف نویز کور در جدول درج شده است که کمی از دقت حذف نویز غیر کور کمتر است. در خصوص DudeNet میزان PSNR مدل حذف نویز کور تفاوت زیادی با مدل غیر کور دارد و PSNR مدل غیر کور از PSNR شبکه عصبی FFDNet و ADNet نیز بیشتر است و اگر از مقادیر PSNR مدل غیر کور برای رسم نمودار استفاده شود، DudeNet بعد از این شبکه‌های عصبی قرار گرفته و نسبت به تمام شبکه‌های عصبی تک شاخه عملکرد بهتری از خود ارائه داده است.

شبکه عصبی N3Net یکی از شبکه‌های عصبی این گروه است که بر روی مجموعه داده BSD68 نتایج بسیار بهتری از مجموعه داده SET12 ارائه کرده است. این شبکه عصبی از وابستگی‌های غیر محلی استفاده می‌کند و احتمالاً به دلیل پیچیدگی و تنوع بیشتر تصاویر مجموعه داده BSD68 این شبکه بهتر توانسته از وابستگی‌های غیر محلی تصاویر استفاده کند و نتایج بهتری را در بازیابی تصاویر این مجموعه داده ارائه کرده و در ردیف شبکه‌های عصبی چند شاخه قرار گرفته است. به عبارت دیگر استفاده از ویژگی‌های غیر محلی بر روی این مجموعه داده به

شبکه‌های عصبی، عملکرد بهتری ارائه کردند. NLRN در تمامی سطوح نویز و MWCNN در پنج سطح نویز از پایگاه داده‌های مختلف، جزء ۵ روش برتر هستند. در حذف نویز تصاویر رنگی با سطوح نویز مختلف شبکه عصبی حذف نویز DRANet [۹۲] بعد از سه شبکه عصبی KBNNet، SCUNet و SwinIR نتایج بهتری نسبت به سایر شبکه‌های عصبی ارائه داده است. بررسی شباهت‌های ساختاری این سه شبکه در بخش ۱-۴ بیان شده است. همانطور که در جدول‌های ۲ تا ۴ نشان داده شده است هر دو شبکه عصبی KBNNet [۸۸]، SCUNet [۴۴] نسبت به سایر روش‌ها بالاترین مقادیر PSNR را ارائه کرده و نتایجی بسیار نزدیک به یکدیگر ارائه کرده‌اند. همانطور که در بخش ۱-۴ به طور کامل شباهت‌های ساختاری شبکه‌های عصبی موفق بررسی شد، KBNNet و SCUNet از مکانیزم توجه بهره می‌برند. در SCUNet توجه مبتنی بر مبدل‌ها یا مکانیزم خود توجه چند وجهی (MSA) استفاده شده است و توجه در KBNNet، ترکیبی از مکانیزم توجه کانال و مکانیزم توجه فضایی است. شباهت این دو مکانیزم توجه در محاسبه وزن‌ها است. در مکانیزم توجه مورد استفاده در این دو شبکه عصبی محاسبه وزن‌ها به صورت جهانی نیست و وزن‌های توجه در یک همسایگی از هر پیکسل محاسبه می‌شود. شبکه عصبی SwinIR [۸۲] نزدیکترین کارایی را به KBNNet و SCUNet دارد. در این شبکه عصبی نیز مشابه شبکه عصبی SCUNet از مکانیزم توجه مبتنی بر مبدل‌ها استفاده شده است و محاسبه وزن‌ها در این مکانیزم توجه، به صورت محلی انجام می‌شود. این شباهت‌های ساختاری ممکن است علت کارایی حذف نویز مشابه این سه شبکه عصبی باشد.

با توجه به تعداد زیاد شبکه‌های عصبی و وجود دیتاست‌های مختلف با مقادیر نویز متفاوت، استفاده از نمودار برای مقایسه شبکه‌های عصبی حذف نویز می‌تواند به درک بهتر تفاوت‌ها و شباهت‌های مدل‌های مختلف و ارزیابی مدل‌ها در شرایط متفاوت کمک کند. برای همین منظور در نمودارهای شکل‌های (۲۷) تا (۳۱) عملکرد روش‌های مختلف حذف نویز بر روی مجموعه داده‌های خاکستری SET12 و BSD68 و مجموعه داده رنگی BSD68 با مقادیر مختلف نویز نشان داده شده است. در شکل (۲۷) نمودار میزان PSNR روش‌های مختلف برای بازیابی مجموعه داده SET12، با شدت‌های مختلف نویز ارائه شده است. در شکل (۲۸) نمودار میزان PSNR روش‌های مختلف حذف نویز برای بازیابی مجموعه داده BSD68 با شدت‌های مختلف نویز نشان داده شده است. در هر نمودار مدل‌ها بر اساس میزان PSNR تصاویر بازیابی شده همان مجموعه داده، با شدت نویز $\sigma=50$ ، مرتب شده است.

در یک مقایسه بصری کلی، با توجه به اینکه در نمودارهای شکل (۲۷) و (۲۸) مدل‌های حذف نویز بر اساس میزان PSNR

¹ Multi Head Self Attention

برخلاف سایر روش‌ها که مشابهت‌های غیر محلی را در تمام تصویر جستجو می‌کنند، در NLRN جستجوی مشابهت‌های غیر محلی در یک همسایگی انجام می‌شود. این موضوع، مشابه مکانیزم توجه سه شبکه عصبی برتر حذف نویز (SwinIR، KBNNet، SCUNet) است و ممکن است دلیل عملکرد خوب NLRN در فرآیند حذف نویز تصویر باشد. در خصوص شبکه‌های عصبی برتر (SwinIR، KBNNet، SCUNet) که بهترین نتایج را بر روی مجموعه‌های داده مختلف و با سطوح نویز مختلف ارائه کرده‌اند، در ابتدای این زیر بخش توضیحات کاملی ارائه شده است. اختلاف مقدار PSNR تصاویر بازیابی شده با شدت نویز $\sigma=50$ با تصاویر تخریب شده با شدت نویز $\sigma=15$ برای دو شبکه عصبی DnCNN و MADRNet نسبت به سایر شبکه‌ها بیشتر است (بر روی هر دو مجموعه داده) و این دو شبکه برای نویزهای با شدت پایین بسیار موفق‌تر از نویزهای با شدت بالا عمل کرده و افزایش شدت نویز عملکرد شبکه را به شدت کاهش داده است.

مقایسه روش‌های حذف نویز بر اساس معیار SSIM در جدول‌های (۵) و (۶) ارائه شده است. SSIM از نظر بینایی بصری انسان، بهتر از PSNR است. سه مقدار برتر در هر شدت نویز با زیر خط مشخص شده است. برای بررسی و مقایسه بهتر مقادیر SSIM، نمودار مقادیر SSIM تصاویر بازیابی شده مجموعه داده SET12 با روش‌های مختلف حذف نویز، برای مقادیر مختلف نویز در شکل (۲۹) ارائه شده است. در شکل (۳۰) نیز نمودار مقادیر SSIM برای تصاویر بازیابی شده مجموعه داده BSD68 با روش‌های مختلف حذف نویز، برای مقادیر مختلف نویز ارائه شده است. در هر نمودار مدل‌ها بر اساس میزان SSIM همان مجموعه داده و $\sigma=50$ مرتب شده است. در این نمودارها نیز همانند نمودارهای شکل‌های (۲۷) و (۲۸) ضعیف‌ترین روش‌ها، روش‌های مبتنی بر مدل هستند.

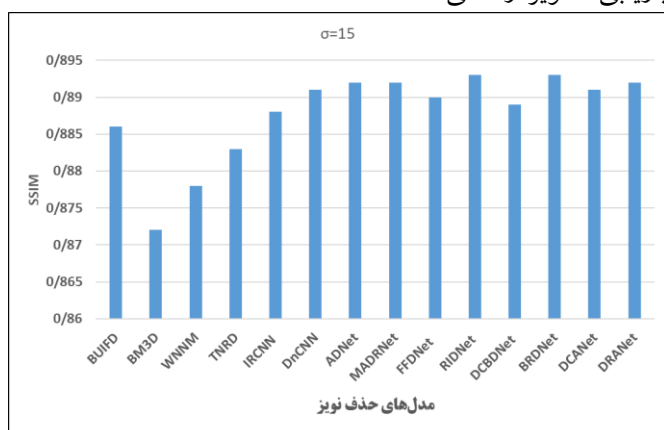
بعد از روش‌های مبتنی بر مدل، شبکه‌های عصبی تک شاخه قرار دارند، سپس شبکه‌های عصبی چند شاخه و بعد از آن شبکه‌های عصبی که در کنار چند شاخه بودن از مکانیزم‌های متعدد توجه استفاده می‌کنند، قرار گرفته‌اند. عملکرد شبکه عصبی BUFD در شدت‌های بالای نویز ضعیف‌تر از روش‌های مبتنی بر مدل است ولی در شدت کم نویز، این شبکه عصبی در ردیف شبکه‌های عصبی تک شاخه DnCNN و IRCNN قرار دارد. (دلیل این موضوع در همین بخش در مقایسه مقادیر PSNR ارائه شده است) بعد از این شبکه‌های عصبی، شبکه‌های عصبی RIDNet، ADNet، MADRNet، FFDNet و DCBDNet قرار دارند. شبکه‌های عصبی RIDNet، ADNet و MADRNet از یک ماژول توجه استفاده می‌کنند. FFDNet یک شبکه عصبی تک شاخه و بدون مکانیزم توجه است، با این حال در میان این ۵ شبکه عصبی و برای شدت نویز زیاد، بالاتر از

اندازه استفاده از شبکه عصبی چند شاخه موثر بوده است. شبکه عصبی FFDNet نیز در مقایسه با سایر شبکه‌های عصبی این گروه، عملکرد بهتری بر روی مجموعه داده BSD68 ارائه داده است. این شبکه عصبی بر روی مجموعه داده BSD68 نتایج بسیار بهتری از مجموعه داده SET12 ارائه کرده است. با توجه به اینکه تنها تفاوت این شبکه با DnCNN، ورودی کنترل FFDNet است که این شبکه را در مقابل سطوح مختلف نویز انعطاف پذیر می‌کند، احتمالاً این ورودی کنترل موجب شده تا این شبکه با سطوح مختلف نویز تطبیق یافته و بهتر توانسته ویژگی‌های بافت تصویر را از الگوهای نویز با سطوح مختلف استخراج نماید و بر روی تصاویر پیچیده‌تر نتایج بهتری ارائه داده است.

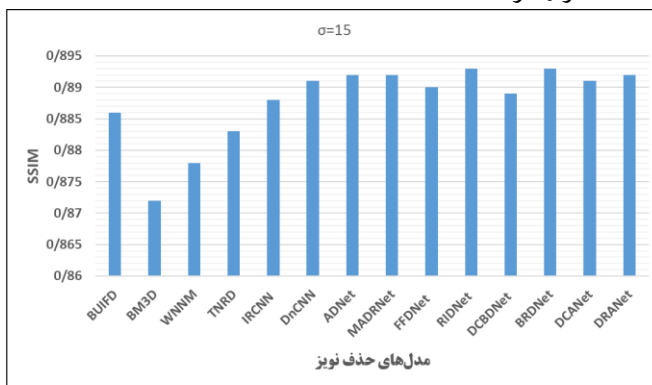
پس از این مدل‌ها، شبکه‌های عصبی RIDNet، MADRNet، BRDNet و DCBDNet قرار دارند که همگی شبکه‌های عصبی چند شاخه هستند و مقدار PSNR آن‌ها نزدیک به هم و از شبکه‌های عصبی قبلی بیشتر است. بنابراین این دو نمودار نیز نشان می‌دهند که استفاده از شبکه‌های عصبی چند شاخه برای حذف نویز موجب بهبود عملکرد شبکه عصبی و افزایش کیفیت تصویر بازیابی شده می‌شود. برخی از این شبکه‌های عصبی مانند RIDNet و MADRNet از یک ماژول توجه نیز استفاده می‌کنند.

بعد از این شبکه‌های عصبی، شبکه‌های عصبی DCANet، RDAN و DRANet قرار دارند. ویژگی‌های مشترک این شبکه‌ها استفاده از چندین ماژول توجه در کنار یکدیگر است. با توجه به اینکه RDAN شبکه عصبی تک شاخه است ترتیب قرارگیری این مدل‌ها نشان می‌دهد که استفاده از ماژول‌های توجه مناسب، در کنار یکدیگر، می‌تواند اثری بیشتر از شبکه‌های عصبی چند شاخه در بهبود مقدار PSNR تصاویر بازیابی شده داشته باشد (در ادامه اشاره می‌شود که در خصوص معیار SSIM عکس این موضوع صحیح است). قرارگیری شبکه عصبی ATDNet در کنار این شبکه‌های عصبی نشان می‌دهد ساختار بلوک‌های رسوبی دروازه‌ای (که در بخش ۳-۴ در خصوص آن توضیح داده شد) به جز اینکه در نوع عملکرد (از نظر اینکه به نقشه‌های ویژگی وزن‌دهی می‌کند) مشابه مکانیزم توجه است، از نظر کارایی نیز، همانند استفاده از ترکیب چند مکانیزم توجه موجب بهبود عملکرد شبکه عصبی ATDNet در حذف نویز تصاویر (بر اساس معیار PSNR) شده است. شبکه عصبی MWCNN نیز با به کارگیری تبدیل موجک دوبعدی و ساختار UNet توانسته در ردیف موفق‌ترین شبکه‌های عصبی حذف نویز قرار گیرد. شبکه عصبی NLRN نیز یک شبکه عصبی است که از مشابهت‌های غیر محلی بهره‌برداری می‌کند. این شبکه عصبی یک شبکه بازگشتی است و توانسته در ردیف بهترین شبکه‌های عصبی حذف نویز قرار گیرد. هرچند روش‌های دیگری مانند NLCNN1 [۶۷] و N3Net [۶۸] نیز از و مشابهت‌های غیر محلی بهره‌برداری کرده‌اند ولی تنها NLRN توانسته عملکردی مشابه بهترین شبکه‌های عصبی حذف نویز ارائه کند.

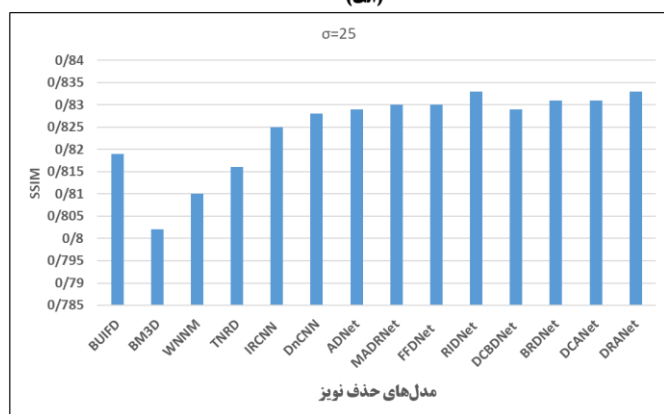
شبکه تخمین نویز باعث شده، این دو شبکه در شدت نویز بالا، عملکرد حذف نویز بهتری از نظر بینایی انسان (معیار SSIM) ارائه کنند. بر خلاف این دو مدل، مدل‌های RIDNet، BRDNet و MADRNet در سطوح نویز پایین نسبت به سطوح نویز بالا عملکرد بسیار بهتری ارائه کرده‌اند. وجه اشتراک این مدل استفاده از کانولوشن توسعه یافته به دلیل افزایش میدان پذیرش (در یک همسایگی) در سطوح نویز پایین اثرات قابل توجهی در بازیابی تصویر دارد ولی در سطوح نویز بالا به دلیل شدت بالای تخریب، استفاده از کانولوشن توسعه یافته، تاثیر کمتری در بازیابی تصویر ارائه می‌کند.



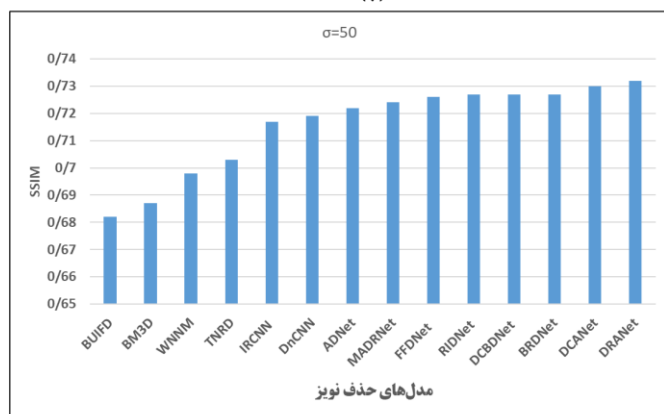
(الف)



(ب)

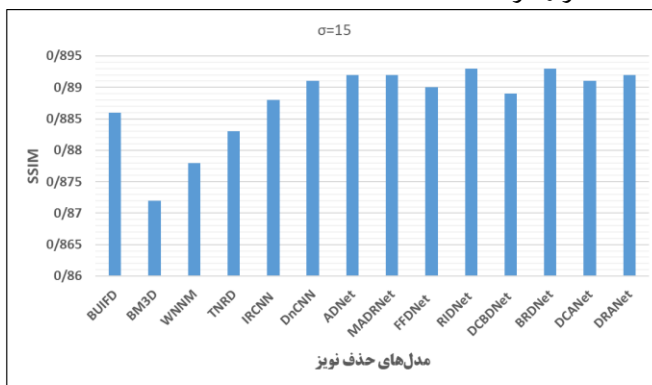


(ج)

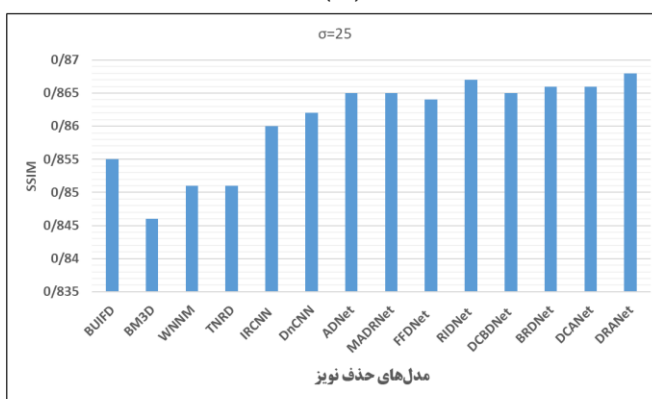


(د)

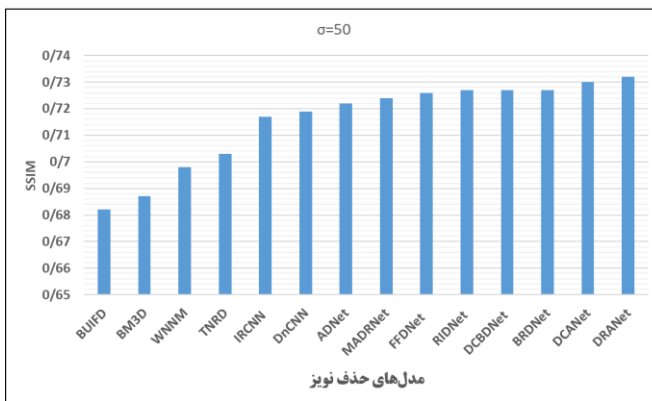
MADRNet و ADNet (در BSD68) که دارای مکانیزم توجه هستند، قرار گرفته است.



(ه)



(و)



(ز)

شکل (۲۹) نمودار ارزیابی روش‌های مختلف حذف نویز بر اساس معیار SSIM برای بازیابی مجموعه داده SET12 با مقادیر مختلف نویز. مدل‌های مختلف حذف نویز بر اساس PSNR تصاویر تخریب شده با شدت نویز $\sigma=50$ مرتب شده است. الف) شدت نویز $\sigma=15$ (ب) شدت نویز $\sigma=25$ (ج) شدت نویز $\sigma=50$

همچنین شبکه عصبی DCBDNet که یک شبکه عصبی چند شاخه بدون مکانیزم توجه است و از نظر کارایی حذف نویز در صدر این پنج شبکه عصبی قرار گرفته است. با توجه به اینکه FFDNet و DCBDNet هر دو از ساختاری تطبیق پذیر با سطوح مختلف نویز بهره می‌برند (FFDNet ورودی کنترلی برای تخمین شدت نویز دارد و DCBDNet از زیر شبکه تخمین نویز بهره می‌برد) و در مقابل سطوح مختلف نویز به صورت تطبیقی عمل می‌کنند، احتمالاً استفاده از ورودی کنترلی شدت نویز و زیر

شکل (۳۰) نمودار ارزیابی روش‌های مختلف حذف نویز بر اساس معیار SSIM برای بازیابی مجموعه داده BSD68 با مقادیر مختلف نویز. مدل‌های مختلف حذف نویز بر اساس PSNR تصاویر

CBSD68 نیز، بعد از شبکه‌های عصبی تک شاخه، شبکه‌های عصبی چند شاخه BRDNet، RIDNet، DudeNet، MADRNet و DCBDNet قرار می‌گیرند و مقدار PSNR آن‌ها از شبکه‌های قبلی بیشتر است. بنابراین این نمودار نیز نشان می‌دهد که استفاده از شبکه‌های عصبی چند شاخه برای حذف نویز تصاویر رنگی، همانند تصاویر خاکستری، موجب بهبود عملکرد شبکه عصبی و افزایش کیفیت تصویر بازیابی شده می‌شود. در میان این شبکه‌ها، شبکه‌های عصبی MADRNet و RIDNet از یک مکانیزم توجه بهره می‌برند. شبکه‌های عصبی BRDNet، MADRNet و RIDNet از کانولوشن توسعه یافته استفاده می‌کنند. این سه شبکه عصبی در سطوح نویز پایین، عملکرد بهتری نسبت به شبکه‌های عصبی DCBDNet، DCANet و ATDNet که از مکانیزم توجه نیز استفاده می‌کنند، ارائه کرده‌اند. عملکرد این سه شبکه عصبی بر روی تصاویر خاکستری است و احتمالاً استفاده از کانولوشن توسعه یافته به دلیل افزایش میدان پذیرش (در یک همسایگی) در سطوح نویز پایین اثرات قابل توجهی در بازیابی تصویر دارد ولی در سطوح نویز بالا به دلیل شدت بالای تخریب، استفاده از کانولوشن توسعه یافته، تاثیر کمتری در بازیابی تصویر ارائه می‌کند. بعد از این شبکه‌های عصبی، شبکه‌های عصبی DCANet، ATDNet، DRANet قرار دارند. ویژگی‌های مشترک این شبکه‌ها (به جز ATDNet) استفاده از چندین ماژول توجه در کنار یکدیگر است. این سه شبکه عصبی در بازیابی تصاویر خاکستری نیز عملکردی مشابه داشتند و در مقایسه با سایر شبکه‌های حذف نویز، عملکرد آن‌ها در بازیابی تصاویر رنگی نیز مشابه عملکردشان در بازیابی تصاویر خاکستری است. همانطور که در ابتدای این بخش (بخش بازیابی تصاویر خاکستری) گفته شد، قرارگیری شبکه عصبی ATDNet در کنار این شبکه‌های عصبی نشان می‌دهد که ساختار بلوک‌های رسوبی دروازه‌ای (که در بخش ۳-۴ در خصوص آن توضیح داده شد) به جز اینکه در نوع عملکرد (از نظر اینکه به نقشه‌های ویژگی وزن دهی می‌کند) مشابه مکانیزم توجه است، از نظر کارایی نیز، همانند استفاده از ترکیب چند مکانیزم توجه، موجب بهبود عملکرد شبکه عصبی ATDNet در حذف نویز (بر اساس معیار PSNR) شده است. در خصوص سایر شبکه‌های عصبی سمت راست نمودار (SCUNet، KBNNet، SwinIR) که بهترین نتایج را بر روی مجموعه‌های داده مختلف و با سطوح نویز مختلف ارائه کرده‌اند، در ابتدای این زیر بخش توضیحات کاملی ارائه شده است.

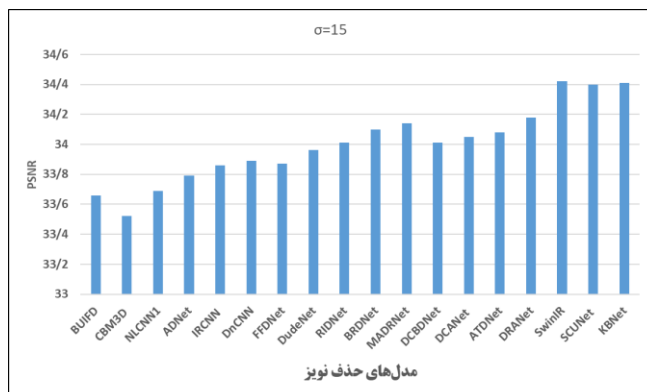
تخریب شده با شدت نویز $\sigma=50$ مرتب شده است. الف) شدت نویز $\sigma=15$ ب) شدت نویز $\sigma=25$ ج) شدت نویز $\sigma=50$

از نظر معماری DCBDNet و BRDNet دارای ساختاری بسیار مشابه هستند و از ترکیب دو شبکه عصبی با ساختاری متفاوت تشکیل شده‌اند و هیچ یک، از مکانیزم توجه استفاده نمی‌کنند. عملکرد بهتر این دو شبکه از شبکه‌های عصبی تک شاخه که از مکانیزم توجه استفاده می‌کنند، نشان دهنده این است که استفاده از شبکه دو شاخه، نسبت به مکانیزم توجه در بازیابی کیفی تر تصاویر از نظر بینایی انسان (معیار SSIM) بسیار موثر عمل نماید (در معیار PSNR عکس این موضوع صحیح بود). شبکه‌های عصبی DCANet و DRANet نسبت به سایر شبکه‌های عصبی عملکرد بسیار خوبی ارائه کرده‌اند. ویژگی‌های مشترک این شبکه‌های عصبی استفاده از شبکه عصبی دو شاخه با ساختار کاملاً متفاوت در کنار استفاده از چندین ماژول توجه در کنار یکدیگر است که موجب بهبود عملکرد شبکه عصبی در حذف نویز تصاویر شده است. اختلاف مقدار SSIM تصاویر بازیابی شده با شدت نویز $\sigma=50$ با تصاویر تخریب شده با شدت نویز $\sigma=15$ برای شبکه عصبی BUIFD نسبت به سایر شبکه‌ها بیشتر است (بر روی هر دو مجموعه داده) و این موضوع بیانگر این است که این شبکه برای نویزهای با شدت پایین بسیار موفق تر از نویزهای با شدت بالا عمل می‌کند و افزایش شدت نویز باعث کاهش شدید عملکرد این شبکه عصبی شده است (در مقایسه معیار PSNR در ابتدای این فصل دلیل این موضوع بررسی شد).

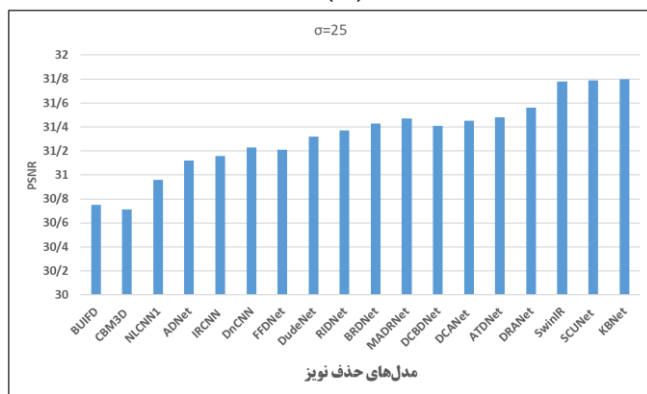
نتایج ارزیابی عملکرد روش‌های مختلف حذف نویز در بازیابی تصاویر رنگی، در جدول (۴) ارائه شده است. با توجه به تنوع مدل‌ها و سطوح مختلف نویز، برای درک بهتر عملکرد مدل‌های حذف نویز و مقایسه آن‌ها با یکدیگر از نمودار استفاده شده است. شکل (۳۱) نمودار مقادیر PSNR تصاویر بازیابی شده مجموعه داده CBSD68 که با سطوح مختلف نویز تخریب شده‌اند را با استفاده از مدل‌های مختلف را نشان می‌دهد. در همه نمودارها مدل‌ها بر اساس میزان PSNR تصاویر بازیابی شده همان مجموعه داده، با شدت نویز $\sigma=50$ ، مرتب شده است. در این نمودار نیز ابتدا روش حذف نویز مبتنی بر مدل معیار CBM3D و شبکه‌های عصبی BUIFD، NLCNN1، ADNet، IRCNN، DnCNN و FFDNet قرار دارند که به جز BUIFD همگی شبکه‌های عصبی تک شاخه هستند. همانطور که در ابتدای این بخش توضیح داده شد، شبکه عصبی BUIFD از نظر معماری از الگوی شبکه‌های چند شاخه پیروی نمی‌کند و صرفاً یک ساختار ثابت در شاخه‌های مختلف تکرار شده است (در قسمت مقایسه نتایج PSNR مجموعه داده BSD68 به طور کامل توضیح داده شده است)، به همین دلیل عملکرد آن در طیف شبکه‌های عصبی تک شاخه قرار می‌گیرد و حتی در شدت نویز بالا، عملکردی ضعیف‌تر از روش‌های مبتنی بر مدل ارائه کرده است. همانند مجموعه‌های داده خاکستری، در مجموعه داده

۵- نتیجه گیری

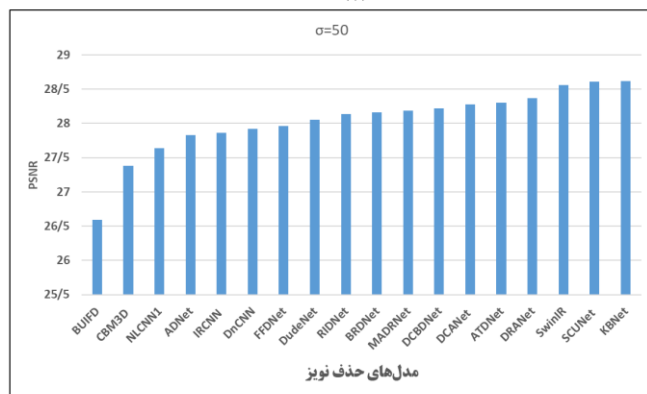
حذف نویز تصویر یک موضوع فعال در پردازش تصویر سطح پایین^۱ است و تاکنون تحقیقات بسیاری در این زمینه انجام شده است. تخریب تصویر به وسیله نویز اجتناب ناپذیر است و عوامل مختلفی در فرآیند دریافت و انتقال تصویر موجب تخریب اطلاعات مفید تصویر می‌شوند؛ بنابراین حذف نویز تصویر یک پیش نیاز ضروری برای تمام وظایف پردازش تصویر است. شبکه‌های عصبی کانولوشن عمیق در سال‌های اخیر نتایج بسیار خوبی را در مسائل مختلف بازیابی تصویر، از جمله حذف نویز تصویر ارائه کرده‌اند و پژوهش‌های بسیاری در این زمینه انجام شده است. در این پژوهش ابتدا پس از بررسی چالش‌ها و ضرورت حذف نویز تصاویر و کاربردهای مختلف آن، روش‌های حذف نویز مبتنی بر مدل و مزایا و معایب آن‌ها مورد بررسی قرار گرفته است؛ سپس روش‌های مبتنی بر یادگیری افتراقی و به ویژه شبکه‌های عصبی حذف نویز، از نظر اجزای تشکیل دهنده و معماری شبکه، مورد بررسی قرار گرفته و ساختارهای تشکیل دهنده این شبکه‌ها مانند اتصالات رسوبی، کانولوشن توسعه‌یافته، شبکه‌های عصبی چند شاخه، بهره‌برداری از مشابهت‌های غیر محلی، مکانیزم توجه و مبدل‌ها که موجب بهبود عملکرد حذف نویز شده است، معرفی شده‌اند. بررسی ساختار شبکه‌های عصبی حذف نویز موفق نشان می‌دهد که استفاده از شبکه‌های عصبی چند شاخه، استفاده از مکانیزم توجه ترکیبی در سطوح مختلف انتزاع (لایه‌های مختلف شبکه عصبی)، به خصوص استفاده از ساختار مبدل‌ها در شبکه‌های عصبی کانولوشن (و همچنین ترکیب تمام این ساختارها با یکدیگر) از مهمترین عوامل بهبود عملکرد شبکه‌های عصبی کانولوشن عمیق حذف نویز است. در بخشی از این مقاله ساختار کلی مشترک شبکه‌های عصبی حذف نویز مدرن، به همراه عملکرد بخش‌های مختلف این ساختار مورد بررسی قرار گرفته است. این ساختار می‌تواند به عنوان مبنایی برای طراحی شبکه‌های عصبی حذف نویز مورد استفاده قرار گیرد. در بخش پایانی مقاله نیز مقایسه جامعی در خصوص عملکرد شبکه‌های عصبی حذف نویز بر اساس معیارهای PSNR و SSIM انجام شده و رابطه بین استفاده از ماژول‌های مختلف و بهبود عملکرد شبکه عصبی حذف نویز ارائه شده است.



(الف)



(ب)



(ج)

شکل (۳۱) نمودار ارزیابی روش‌های مختلف حذف نویز بر اساس معیار PSNR برای بازیابی مجموعه داده CBS68 (تصاویر رنگی) با مقادیر مختلف نویز. مدل‌های مختلف حذف نویز بر اساس PSNR تصاویر تخریب شده با شدت نویز $\sigma=50$ مرتب شده است. (الف) شدت نویز $\sigma=15$ (ب) شدت نویز $\sigma=25$ (ج) شدت نویز $\sigma=50$

جدول (۲) میانگین PSNR تصاویر بازیابی شده مجموعه داده SET12 با سطوح مختلف نویز

$\sigma=75$	$\sigma=50$	$\sigma=25$	$\sigma=15$	روش حذف نویز
۲۴/۹۱	۲۶/۷۲	۲۹/۹۷	۳۲/۳۷	BM3D [۳۹]
۲۵/۲۳	۲۷/۰۵	۳۰/۲۶	۳۲/۷۰	WNNM [۴۲]
۲۵/۰۷	۲۶/۷۸	۳۰/۰۳	-	MLP [۵۲]
-	۲۶/۸۱	۳۰/۰۶	۳۲/۵۰	TNRD [۴۴]
۲۵/۲۰	۲۷/۱۸	۳۰/۴۳	۳۲/۸۶	DnCNN [۵]
۲۵/۴۹	۲۷/۳۲	۳۰/۴۳	۳۲/۷۵	FFDNet [۲۰]
-	۲۷/۳۷	۳۰/۵۲	-	SSDA [۵۸]
-	۲۷/۷۴	۳۰/۷۹	۳۳/۱۵	MWCNN [۶۳]
۲۵/۹۰	۲۷/۴۳	۳۰/۵۵	-	N3Net [۶۸]
-	۲۷/۶۴	۳۰/۸۰	۳۳/۱۶	NLRN [۶۹]
-	۲۷/۶۰	۳۰/۷۸	۳۳/۱۴	GCN1 [۷۲]
-	۲۷/۴۵	۳۰/۶۱	۳۳/۰۳	BRDNet [۷۷]
-	۲۷/۹۱	۳۱/۰۱	۳۳/۳۶	SwinIR [۱۰۶]
-	۲۷/۶۳	۳۰/۷۷	۳۳/۰۷	ATDNet [۸۰]
-	۲۷/۲۲	۳۰/۳۹	۳۲/۷۷	DudeNet [۸۷]
-	۲۷/۴۶	۳۰/۵۱	۳۲/۷۵	DCBDNet [۵۹]
-	۲۸/۰۴	۳۱/۰۹	۳۳/۴۳	SCUNet [۶۴]
-	۲۷/۵۶	۳۰/۷۰	۳۳/۰۴	RDAN [۱۱۴]
-	۲۷/۴۳	۳۰/۶۰	۳۲/۹۱	RIDNet [۹۷]
-	۲۷/۳۳	۳۰/۴۶	۳۲/۷۷	ADNet [۵۵]
-	۲۷/۶۲	۳۰/۶۹	۳۳/۰۰	DRANet [۹۲]
-	۲۷/۴۴	۳۰/۶۵	۳۳/۱۴	MADRNet [۱۱۱]
-	۲۷/۵۴	۳۰/۵۷	۳۲/۸۳	DCANet [۹۵]
-	۲۸/۰۴	۳۱/۰۸	۳۳/۴۰	KBNet [۱۱۲]

جدول (۳) میانگین PSNR تصاویر بازیابی شده مجموعه داده BSD68 با سطوح مختلف نویز

$\sigma=75$	$\sigma=50$	$\sigma=25$	$\sigma=15$	روش حذف نویز
۲۴/۲۱	۲۵/۶۲	۲۸/۵۷	۳۱/۰۷	BM3D [۳۹]
۲۴/۴۰	۲۵/۸۷	۲۸/۸۳	۳۱/۳۷	WNNM [۴۲]
۲۴/۵۹	۲۶/۰۳	۲۸/۹۶	-	MLP [۵۲]
-	۲۶/۱۹	۲۹/۱۵	۳۱/۶۳	IRCNN [۵۶]
-	۲۵/۹۷	۲۸/۹۲	۳۱/۴۲	TNRD [۴۴]
۲۴/۶۴	۲۶/۲۳	۲۹/۲۳	۳۱/۷۲	DnCNN [۵]
۲۴/۷۹	۲۶/۲۹	۲۹/۱۹	۳۱/۶۳	FFDNet [۲۰]
-	۲۶/۵۳	۲۹/۴۱	۳۱/۸۶	MWCNN [۶۳]
-	۲۶/۰۷	۲۹/۰۳	۳۱/۵۲	NLCNN1 [۶۷]
۲۵/۱۴	۲۶/۳۹	۲۹/۳۰	-	N3Net [۶۸]
-	۲۶/۴۷	۲۹/۴۱	۳۱/۸۸	NLRN [۶۹]
-	۲۶/۳۸	۲۹/۳۵	۳۱/۸۳	GCN1 [۷۳]
۲۲/۶۶	۲۵/۱۱	۲۸/۷۴	۳۱/۳۵	BUIFD [۷۹]
-	۲۶/۳۶	۲۹/۲۹	۳۱/۷۹	BRDNet [۷۷]
-	۲۶/۵۸	۲۹/۵۰	۳۱/۹۷	SwinIR [۱۰۶]
-	۲۶/۴۵	۲۹/۳۹	۳۱/۸۲	ATDNet [۸۰]
-	۲۶/۲۵	۲۹/۱۹	۳۱/۶۴	DudeNet [۸۷]
-	۲۶/۲۸	۲۹/۲۴	۳۱/۲۹	CFSNet [۸۵]
۲۴/۸۶	۲۶/۳۷	۲۹/۲۴	۳۱/۶۵	DCBDNet [۵۹]
-	۲۶/۶۷	۲۹/۵۵	۳۱/۹۹	SCUNet [۶۴]
-	۲۶/۴۶	۲۹/۴۱	۳۱/۸۵	RDAN [۱۱۴]
-	۲۶/۴۰	۲۹/۳۴	۳۱/۸۱	RIDNet [۹۷]
-	۲۶/۲۴	۲۹/۱۴	۳۱/۵۶	ADNet [۵۵]
-	۲۶/۴۷	۲۹/۳۶	۳۱/۷۹	DRANet [۹۲]
-	۲۶/۳۷	۲۹/۳۳	۳۱/۹۴	MADRNet [۱۱۱]
-	۲۶/۴۴	۲۹/۳۱	۳۱/۷۱	DCANet [۹۵]
-	۲۶/۶۵	۲۹/۵۴	۳۱/۹۸	KBNet [۱۱۲]

جدول (۴) میانگین PSNR تصاویر بازیابی شده مجموعه داده CBSD68 با سطوح مختلف نویز

$\sigma=75$	$\sigma=50$	$\sigma=35$	$\sigma=25$	$\sigma=15$	روش حذف نویز
۲۵/۷۴	۲۷/۳۸	۲۸/۸۹	۳۰/۷۱	۳۳/۵۲	CBM3D [۳۹]
-	۲۷/۸۶	۲۹/۵۰	۳۱/۱۶	۳۳/۸۶	IRCNN [۵۶]
۲۴/۴۷	۲۷/۹۲	۲۹/۵۸	۳۱/۲۳	۳۳/۸۹	DnCNN [۵]
۲۶/۲۴	۲۷/۹۶	۲۹/۵۸	۳۱/۲۱	۳۳/۸۷	FFDNet [۲۰]
-	۲۷/۶۴	-	۳۰/۹۶	۳۳/۶۹	NLCNN1 [۶۷]
۲۳/۶۲	۲۶/۵۹	۲۸/۸۱	۳۰/۷۵	۳۳/۶۶	BUIFD [۷۹]
-	۲۸/۱۴	-	۳۱/۳۷	۳۴/۰۱	RIDNet [۹۷]
۲۶/۴۳	۲۸/۱۶	۲۹/۷۷	۳۱/۴۳	۳۴/۱۰	BRDNet [۷۷]
-	۲۸/۵۶	-	۳۱/۷۸	۳۴/۴۲	SwinIR [۱۰۶]
-	۲۸/۳۰	-	۳۱/۴۸	۳۴/۰۸	ATDNet [۸۰]
-	۲۸/۰۶	-	-	۳۳/۹۷	CResMD [۸۱]
-	۲۸/۳۱	-	-	۳۶/۳۳	DUBD [۸۴]
-	۲۸/۰۵	۲۹/۶۹	۳۱/۳۲	۳۳/۹۶	DudeNet [۸۷]
۲۶/۵۴	۲۸/۲۲	۲۹/۸۱	۳۱/۴۱	۳۴/۰۱	DCBDNet [۵۹]
-	۲۸/۶۱	-	۳۱/۷۹	۳۴/۴۰	SCUNet [۶۴]
-	۲۷/۸۳	۲۹/۴۸	۳۱/۱۲	۳۳/۷۹	ADNet [۵۵]
-	۲۸/۳۷	-	۳۱/۵۶	۳۴/۱۸	DRANet [۹۲]
-	۲۸/۱۹	۲۹/۷۸	۳۱/۴۷	۳۴/۱۴	MADRNet [۱۱۱]
۲۶/۶۱	۲۸/۲۸	۲۹/۸۶	۳۱/۴۵	۳۴/۰۵	DCANet [۹۵]
-	۲۸/۶۲	-	۳۱/۸۰	۳۴/۴۱	KBNet [۱۱۲]

جدول (۵) مقدار SSIM تصاویر بازیابی شده مجموعه داده SET12 با سطوح مختلف نویز

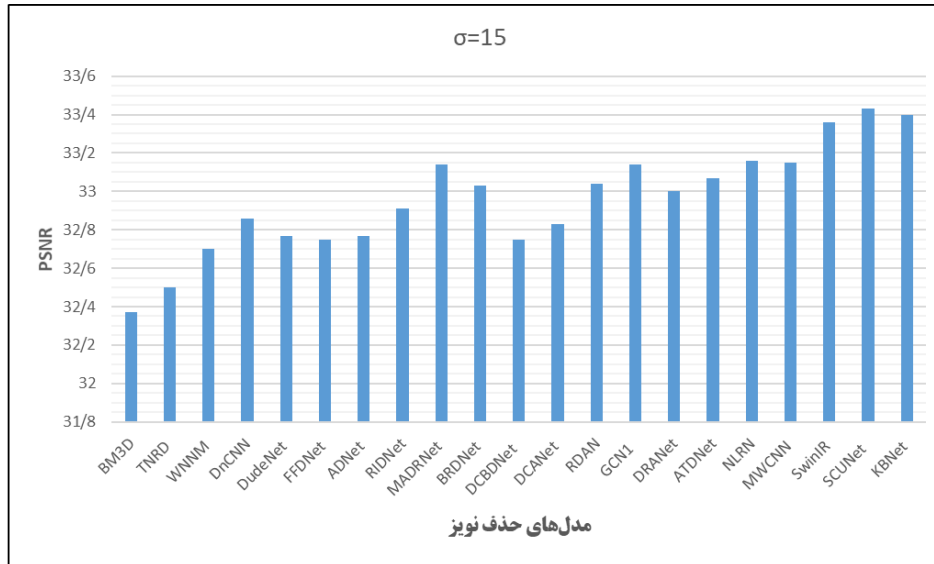
$\sigma=50$	$\sigma=25$	$\sigma=15$	روش حذف نویز
۰/۷۶۶	۰/۸۵۱	۰/۸۹۶	BM3D [۳۹]
۰/۷۵۶	۰/۸۴۶	۰/۸۹۴	WNNM [۴۲]
۰/۷۸۰	۰/۸۶۰	۰/۹۰۱	IRCNN [۵۶]
۰/۷۶۸	۰/۸۵۱	۰/۸۹۶	TNRD [۴۴]
۰/۷۸۳	۰/۸۶۲	۰/۹۰۳	DnCNN [۵]
۰/۷۹۱	۰/۸۶۴	۰/۹۰۳	FFDNet [۲۰]
۰/۷۵۵	۰/۸۵۵	۰/۸۹۹	BUIFD [۷۹]
۰/۷۹۴	۰/۸۶۶	۰/۹۰۶	BRDNet [۷۷]
۰/۷۹۴	۰/۸۶۵	۰/۹۰۲	DCBDNet [۵۹]
۰/۷۹۳	۰/۸۶۷	۰/۹۰۶	RIDNet [۹۷]
۰/۷۹۱	۰/۸۶۵	۰/۹۰۵	ADNet [۵۵]
۰/۸۰۰	۰/۸۶۸	۰/۹۰۶	DRANet [۹۲]
۰/۷۸۹	۰/۸۶۵	۰/۹۰۵	MADRNet [۱۱۱]
۰/۷۹۸	۰/۸۶۶	۰/۹۰۳	DCANet [۹۵]

جدول (۱) مقایسه معماری و ویژگی‌های مختلف شبکه‌های عصبی حذف نویز

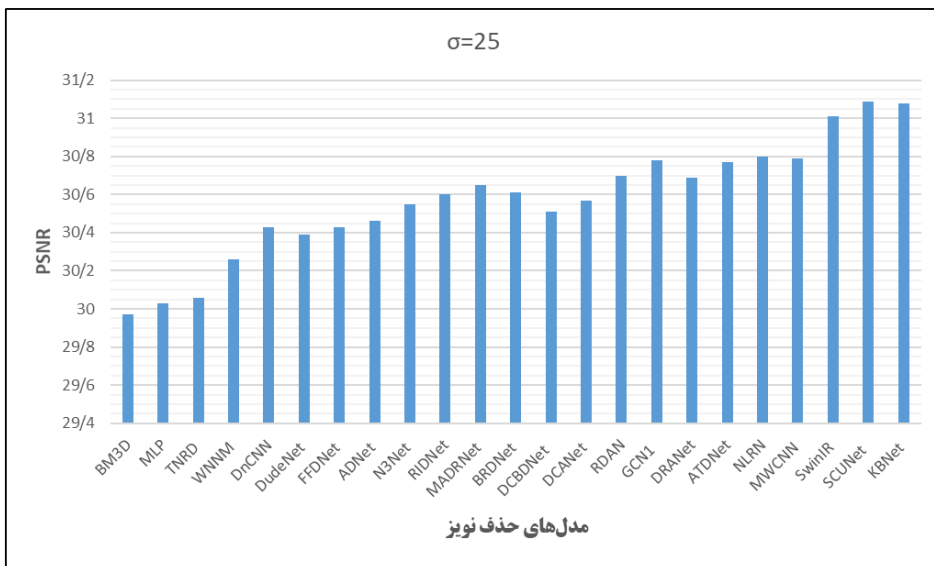
ردیف	عنوان شبکه عصبی حذف نویز	حذف نویز کور	حذف نویز کنترل میزان حذف یافته	کانولوشن توسعه یافته	استفاده از زیر شبکه تخمین نویز	عملکرد چند منظوره	پهرو برداری مستقیم از ویژگی‌های غیر محلی	شبکه عصبی چند شاخه	استفاده از ورودی مختلف برای شاخه‌های مختلف	مبدا/فضایی/کانال/سفارشی (مبتنی بر مکانیزم توجه رسوبی)	رویکرد یادگیری
۱	MLP [۵۲]	-	-	-	-	-	-	-	-	-	-
۲	IRCNN [۵۶]	-	-	✓	-	✓	-	-	-	-	✓
۳	DnCNN [۵]	✓	-	-	-	✓	-	-	-	-	✓
۴	FFDNet [۲۰]	✓*	✓	-	-	-	-	-	-	-	-
۵	SSDA [۵۸]	✓	-	-	-	✓	-	-	-	-	-
۶	RED-Net [۶۰]	✓	-	-	-	✓	-	-	-	-	✓
۷	CBDNet [۶۱]	✓	-	-	✓	-	-	-	-	-	✓
۸	MWCNN [۶۲]	✓	-	ضمنی**	-	✓	-	-	-	-	-
۹	NLCNN1[۶۷]	✓	-	-	-	-	✓	-	-	-	-
۱۰	N3Net [۶۸]	-	-	-	-	✓	✓	-	-	-	✓
۱۱	NLRN [۶۹]	-	-	-	-	✓	✓	-	-	-	✓
۱۲	CPNet [۷۲]	✓*	✓	-	-	-	✓	✓	-	-	✓
۱۳	GCN1 [۷۳]	-	✓	-	-	-	✓	-	-	-	✓
۱۴	BRDNet [۷۷]	✓	-	✓	-	-	-	✓	-	-	✓
۱۵	BUIFD [۷۹]	✓	-	-	✓	-	-	✓	-	-	-
۱۶	ATDNet [۸۰]	✓	-	-	✓	-	-	✓	-	-	✓
۱۷	CResMD [۸۱]	✓*	✓	-	-	✓	-	✓	✓	-	✓
۱۸	AirNet [۸۲]	✓	-	-	ضمنی	✓	-	✓	-	-	-
۱۹	U-WADN [۸۳]	✓	-	-	ضمنی	✓	-	✓	-	-	-
۲۰	DUBD [۸۴]	✓	-	-	ضمنی	-	-	✓	-	-	✓
۲۱	DudeNet [۸۷]	✓	-	✓	-	-	-	✓	-	-	✓
۲۲	CFSNet [۸۵]	✓*	✓	-	-	✓	-	✓	✓	-	-
۲۳	Path-Restore [۸۸]	✓	-	-	✓	-	-	✓	-	-	-
۲۴	DCBDNet [۵۹]	✓	-	✓	✓	-	-	✓	-	-	-
۲۵	SwinIR [۱۰۶]	-	-	-	✓	-	با استفاده از توجه	-	-	استفاده از مبدل‌ها	✓
۲۶	SCUNet [۶۴]	-	-	-	-	-	با استفاده از توجه	-	-	استفاده از مبدل‌ها	✓
۲۷	MAN [۱۰۷]	-	-	-	✓	-	-	-	-	کانال	-
۲۸	RNAN [۱۰۸]	-	-	-	✓	✓	✓	-	-	فضایی و کانال	✓
۲۹	RDAN [۱۱۴]	✓	-	✓	-	-	-	✓	-	فضایی و کانال	✓
۳۰	RIDNet [۹۷]	✓	-	✓	-	-	-	✓	-	کانال	✓
۳۱	ADNet [۵۵]	✓	-	✓	-	-	-	✓	-	کانال	✓
۳۲	DRANet [۹۲]	✓	-	✓	-	-	-	✓	-	فضایی و کانال	✓
۳۳	MADRNet [۱۱۱]	✓	-	✓	ضمنی	-	-	✓	-	سفارشی	✓
۳۴	KBNet [۱۱۲]	✓	-	-	✓	-	-	✓	-	فضایی (سفارشی) و کانال	-
۳۵	DCANet [۹۵]	✓	-	✓	✓	-	-	✓	-	فضایی و کانال	✓

* با دریافت نقشه سطح نویز یا مقدار کنترلی از ورودی

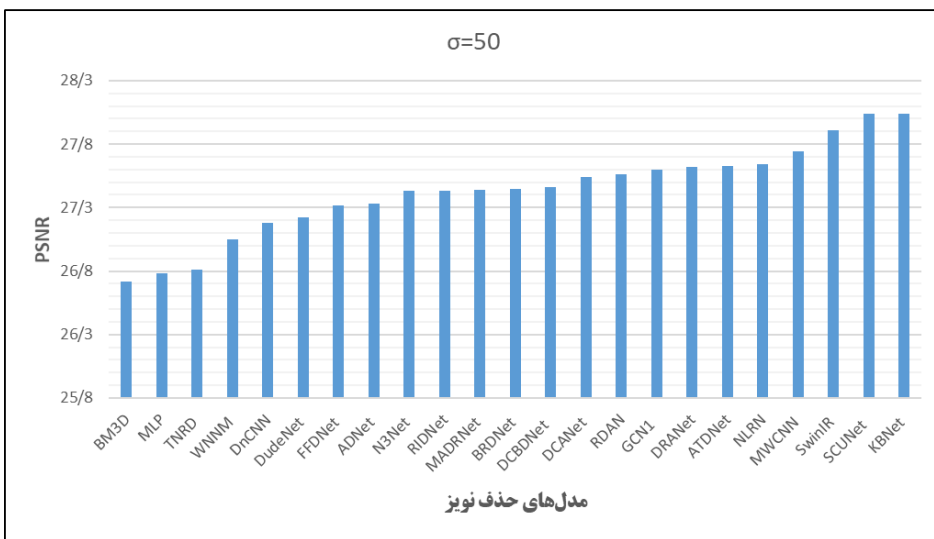
** با استفاده از تبدیل موجک گسسته دو بعدی



(الف)

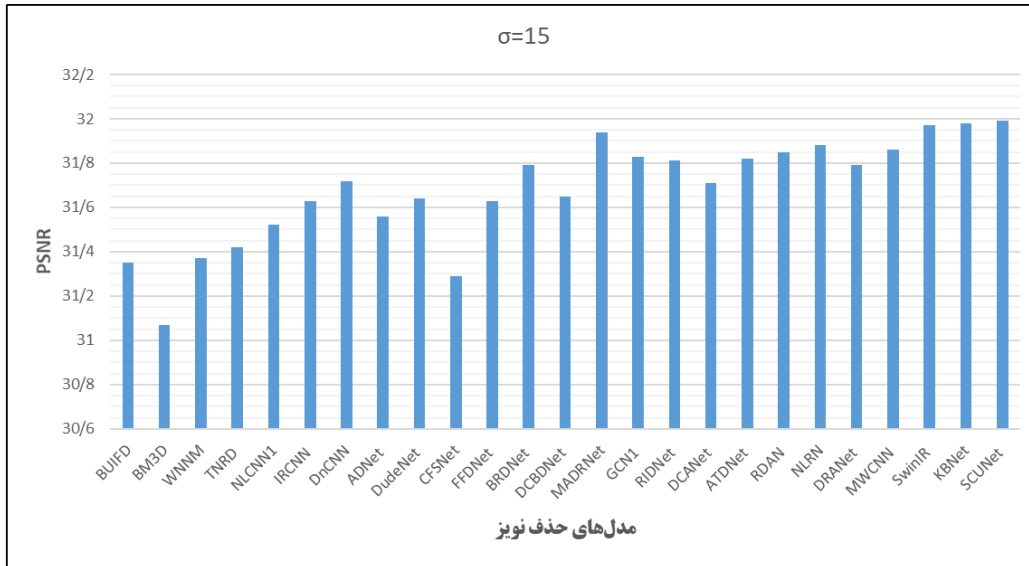


(ب)

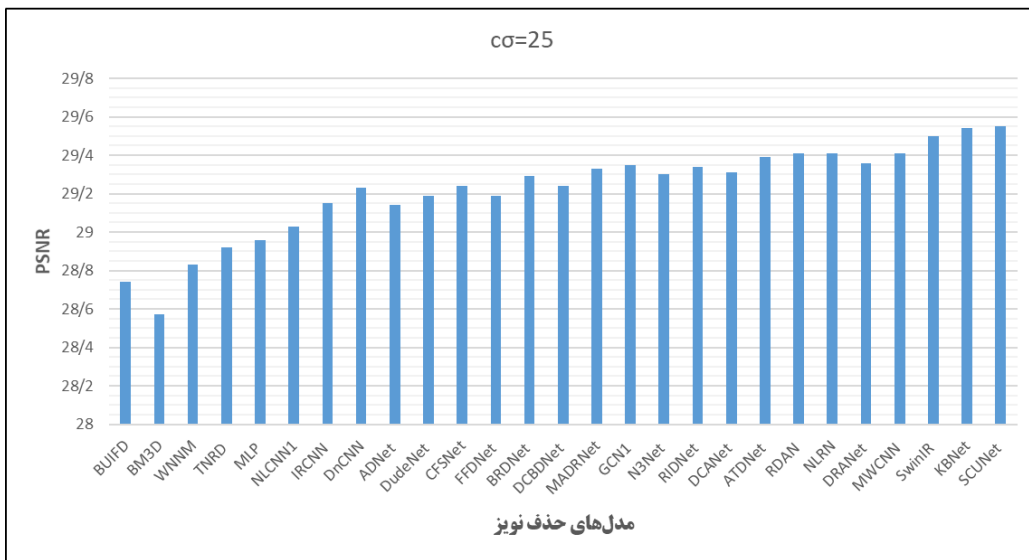


(ج)

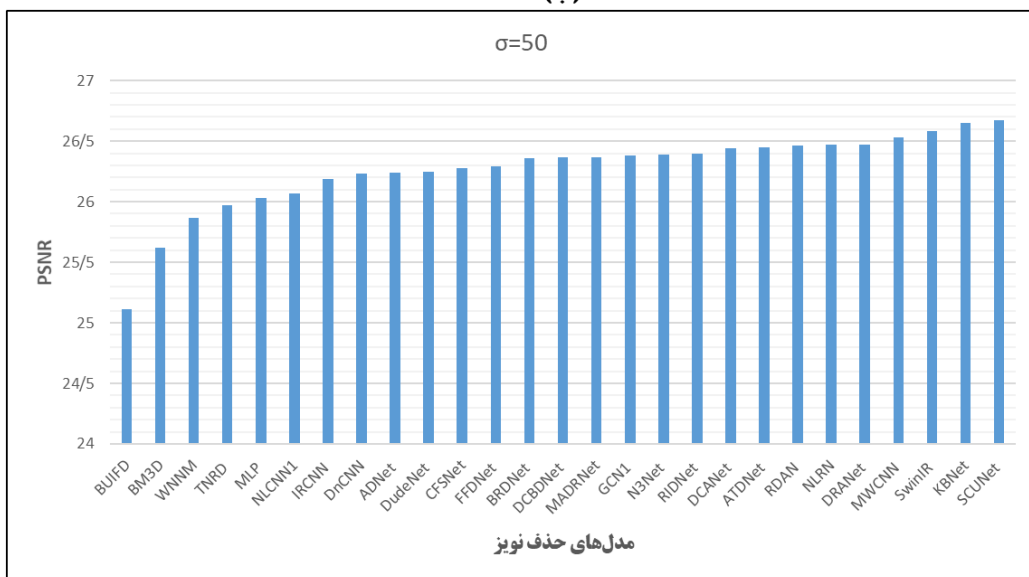
شکل (۲۷) نمودار ارزیابی روش‌های مختلف حذف نویز بر اساس معیار PSNR برای بازیابی مجموعه داده SET12 با مقادیر مختلف نویز. مدل‌های مختلف حذف نویز بر اساس PSNR تصاویر تخریب شده با شدت نویز $\sigma=50$ مرتب شده است. (الف) شدت نویز $\sigma=15$ (ب) شدت نویز $\sigma=25$ (ج) شدت نویز $\sigma=50$



(الف)



(ب)



شکل (۲۸) نمودار ارزیابی روش‌های مختلف حذف نویز بر اساس معیار PSNR برای بازیابی مجموعه داده BSD68 با مقادیر مختلف نویز. مدل‌های مختلف حذف نویز بر اساس PSNR تصاویر تخریب شده با شدت نویز $\sigma=50$ مرتب شده است. (الف) شدت نویز $\sigma=15$ (ب) شدت نویز $\sigma=25$ (ج) شدت نویز $\sigma=50$

[11] Alnuaimy, A.N.H., et al. "BM3D Denoising Algorithms for Medical Image". in *2024 35th Conference of Open Innovations Association (FRUCT)*. 2024.

[12] L, M. and M.A. Dorairangasamy, "An Extensive Review of Significant Researches on Medical Image Denoising Techniques". *International Journal of Computer Applications*, 2013. 64(14): p. 1-12.

[13] Ding, Q., et al., "Statistical Image Reconstruction Using Mixed Poisson-Gaussian Noise Model for X-Ray CT". 2018.

[14] P, P. and V.V. R. "A Comprehensive Analysis of Satellite Image Denoising using Earth Observation Data". in *2023 7th International Conference on Computing Methodologies and Communication (ICCMC)*. 2023.

[15] Ayanala, B., et al., "Satellite Image Denoising Using Machine Learning". 2024.

[16] Wei, S., et al., "Research on SAR image denoising method based on feature extraction". *Electronics Letters*, 2024. 60(8): p. e13180.

[17] "RETRACTED ARTICLE: Image denoising and despeckling methods for SAR images to improve image enhancement performance: a survey". *Multimedia Tools Appl.*, 2021. 80(17): p. 26547-26569.

[18] Liu, S., et al., "LG-DBNet: Local and Global Dual-Branch Network for SAR Image Denoising". *IEEE Transactions on Geoscience and Remote Sensing*, 2024. 62: p. 1-15.

[19] Khasanova, A., A. Makhmutova, and I. Anikin. "Image Denoising for Video Surveillance Cameras Based on Deep Learning Techniques". in *2021 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)*. 2021.

[20] Zhang, K., W. Zuo, and L. Zhang, "FFDNet: Toward a Fast and Flexible Solution for CNN based Image Denoising". *IEEE Transactions on Image Processing*, 2017. PP.

[21] Fan, L., et al., "Brief review of image denoising techniques". *Visual Computing for Industry, Biomedicine, and Art*, 2019. 2(1): p. 7.

[22] Bled, C. and F. Pitié, "Lightweight Video Denoising Using a Classic Bayesian Backbone". 2024.

[23] Maggioni, M., et al. "Efficient Multi-Stage Video Denoising with Recurrent Spatio-Temporal Fusion". in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2021.

[24] Yue, H., et al. "Supervised Raw Video Denoising With a Benchmark Dataset on Dynamic Scenes". in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2020.

[25] Tassano, M., J. Delon, and T. Veit. "FastDVDnet: Towards Real-Time Deep Video Denoising Without Flow Estimation". in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2020.

[26] Vaksman, G., M. Elad, and P. Milanfar. "Patch Craft: Video Denoising by Deep Modeling and Patch Matching". in

جدول (۶) مقدار SSIM تصاویر بازیابی شده مجموعه داده BSD68 با

سطوح مختلف نویز

$\sigma=50$	$\sigma=25$	$\sigma=15$	روش حذف نویز
۰/۶۸۷	۰/۸۰۲	۰/۸۷۲	BM3D [۳۹]
۰/۶۹۸	۰/۸۱۰	۰/۸۷۸	WNNM [۴۲]
۰/۷۱۷	۰/۸۲۵	۰/۸۸۸	IRCNN [۵۶]
۰/۷۰۳	۰/۸۱۶	۰/۸۸۳	TNRD [۴۴]
۰/۷۱۹	۰/۸۲۸	۰/۸۹۱	DnCNN [۵]
۰/۷۲۶	۰/۸۳۰	۰/۸۹۰	FFDNet [۲۰]
۰/۶۸۲	۰/۸۱۹	۰/۸۸۶	BUIFD [۷۹]
۰/۷۲۷	۰/۸۳۱	۰/۸۹۳	BRDNet [۷۷]
۰/۷۲۷	۰/۸۲۹	۰/۸۸۹	DCBDNet [۵۹]
۰/۷۲۷	۰/۸۳۳	۰/۸۹۳	RIDNet [۹۷]
۰/۷۲۲	۰/۸۲۹	۰/۸۹۲	ADNet [۵۵]
۰/۷۳۲	۰/۸۳۳	۰/۸۹۲	DRANet [۹۲]
۰/۷۲۴	۰/۸۳۰	۰/۸۹۲	MADRNet [۱۱۱]
۰/۷۳۰	۰/۸۳۱	۰/۸۹۱	DCANet [۹۵]

مراجع

[1] Bajaj, K., D.K. Singh, and M.A. Ansari, "Autoencoders Based Deep Learner for Image Denoising". *Procedia Computer Science*, 2020. 171: p. 1535-1541.

[2] Joshua, O., S. Ibiyemi, and B. Adu, "A Comprehensive Review On Various Types of Noise in Image Processing". *International Journal of Scientific and Engineering Research*, 2019. 10: p. 388-393.

[3] Boncelet, C., Chapter 7 - Image Noise Models, in *The Essential Guide to Image Processing*, A. Bovik, Editor. 2009, Academic Press: Boston. p. 143-167.

[4] Goyal, B., et al., "Image denoising review: From classical to state-of-the-art approaches". *Information Fusion*, 2020. 55: p. 220-244.

[5] Zhang, K., et al., "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising". *IEEE Transactions on Image Processing*, 2017. 26(7): p. 3142-3155.

[6] Chatterjee, P. and P. Milanfar, "Is Denoising Dead?". *IEEE Transactions on Image Processing*, 2010. 19(4): p. 895-911.

[7] Mohd Sagheer, S.V. and S.N. George, "A review on medical image denoising algorithms". *Biomedical Signal Processing and Control*, 2020. 61: p. 102036.

[8] Tian, J. and L. Chen, "Image despeckling using a non-parametric statistical model of wavelet coefficients". *Biomedical Signal Processing and Control*, 2011. 6(4): p. 432-437.

[9] Zhang, J., et al., "Wavelet and fast bilateral filter based despeckling method for medical ultrasound images". *Biomedical Signal Processing and Control*, 2015. 18: p. 1-10.

[10] Aja-Fernandez, S., C. Alberola-Lopez, and C.F. Westin, "Noise and Signal Estimation in Magnitude MRI and Rician Distributed Images: A LMMSE Approach". *IEEE Transactions on Image Processing*, 2008. 17(8): p. 1383-1398.

- [42] Gu, S., et al. "Weighted Nuclear Norm Minimization with Application to Image Denoising". in *2014 IEEE Conference on Computer Vision and Pattern Recognition*. 2014.
- [43] Dong, W., et al., "Nonlocally Centralized Sparse Representation for Image Restoration". *IEEE Transactions on Image Processing*, 2013. 22(4): p. 1620–1630.
- [44] Chen, Y. and T. Pock, "Trainable Nonlinear Reaction Diffusion: A Flexible Framework for Fast and Effective Image Restoration". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017. 39(6): p. 1256–1272.
- [45] Perona, P. and J. Malik, "Scale-space and edge detection using anisotropic diffusion". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1990. 12(7): p. 629–639.
- [46] Weickert, J., "Anisotropic Diffusion in Image Processing". 1998.
- [47] Schmidt, U. and S. Roth. "Shrinkage Fields for Effective Image Restoration". in *2014 IEEE Conference on Computer Vision and Pattern Recognition*. 2014.
- [48] Roth, S. and M.J. Black. "Fields of Experts: a framework for learning image priors". in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*. 2005.
- [49] Krizhevsky, A., I. Sutskever, and G.E. Hinton, "ImageNet classification with deep convolutional neural networks". *Commun. ACM*, 2017. 60(6): p. 84–90.
- [50] Ioffe, S. and C. Szegedy, Batch normalization: accelerating deep network training by reducing internal covariate shift, in *Proceedings of the 32nd International Conference on International Conference on Machine Learning – Volume 37*. 2015, JMLR.org: Lille, France. p. 448–456.
- [51] He, K., et al. "Deep Residual Learning for Image Recognition". in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016.
- [52] Burger, H.C., C.J. Schuler, and S. Harmeling. "Image denoising: Can plain neural networks compete with BM3D?". in *2012 IEEE Conference on Computer Vision and Pattern Recognition*. 2012.
- [53] Jain, V. and H.S. Seung, Natural image denoising with convolutional networks, in *Proceedings of the 21st International Conference on Neural Information Processing Systems*. 2008, Curran Associates Inc.: Vancouver, British Columbia, Canada. p. 769–776.
- [54] Yu, F. and V. Koltun, *Multi-Scale Context Aggregation by Dilated Convolutions*. 2016.
- [55] Tian, C., et al., "Attention-guided CNN for image denoising". *Neural Networks*, 2020. 124: p. 117–129.
- [56] Zhang, K., et al. "Learning Deep CNN Denoiser Prior for Image Restoration". in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017.
- [57] Larrue, T., X. Meng, and C. Han. "Denoising Videos with Convolutional Autoencoders A Comparison of Autoencoder Architectures". 2018.
- 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*. 2021.
- [27] Liang, J., et al., "VRT: A Video Restoration Transformer". *IEEE Transactions on Image Processing*, 2024. 33: p. 2171–2182.
- [28] Fu, Z., et al., Temporal As a Plugin: Unsupervised Video Denoising with Pre-trained Image Denoisers, in *Computer Vision – ECCV 2024: 18th European Conference, Milan, Italy, September 29–October 4, 2024, Proceedings, Part LVI*. 2024, Springer-Verlag: Milan, Italy. p. 349–367.
- [29] Gonzalez, R.C. and R.E. Woods, *Digital Image Processing (3rd Edition)*. 2006: Prentice-Hall, Inc.
- [30] Al-Ameen, Z., S. Al-Ameen, and G. Sulong, "Latest Methods of Image Enhancement and Restoration for Computed Tomography: A Concise Review". *Applied Medical Informatics*, 2015. 36: p. 1–12.
- [31] Benesty, J., J. Chen, and Y. Huang. "Study of the widely linear Wiener filter for noise reduction". in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*. 2010.
- [32] Pitas, I. and A.N. Venetsanopoulos. "Nonlinear Digital Filters – Principles and Applications". in *The Springer International Series in Engineering and Computer Science*. 1990.
- [33] Ruikang, Y., et al., "Optimal weighted median filtering under structural constraints". *IEEE Transactions on Signal Processing*, 1995. 43(3): p. 591–604.
- [34] Tomasi, C. and R. Manduchi. "Bilateral filtering for gray and color images". in *Sixth International Conference on Computer Vision (IEEE Cat. No. 98CH36271)*. 1998.
- [35] Xie, Q., et al., "Color and direction-invariant nonlocal self-similarity prior and its application to color image denoising". *Science China Information Sciences*, 2020. 63(12): p. 222101.
- [36] Buades, A., B. Coll, and J.-M. Morel, "Nonlocal Image and Movie Denoising". *International Journal of Computer Vision*, 2008. 76(2): p. 123–139.
- [37] Buades, A., B. Coll, and J.M. Morel. "A non-local algorithm for image denoising". in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*. 2005.
- [38] Mairal, J., et al. "Non-local sparse models for image restoration". in *2009 IEEE 12th International Conference on Computer Vision*. 2009.
- [39] Dabov, K., et al., "Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering". *IEEE Transactions on Image Processing*, 2007. 16(8): p. 2080–2095.
- [40] Mairal, J., M. Elad, and G. Sapiro, "Sparse Representation for Color Image Restoration". *IEEE Transactions on Image Processing*, 2008. 17(1): p. 53–69.
- [41] Elad, M. and M. Aharon, "Image Denoising Via Sparse and Redundant Representations Over Learned Dictionaries". *IEEE Transactions on Image Processing*, 2006. 15(12): p. 3736–3745.

- [72] Li, Y., X. Fu, and Z.J. Zha. "Cross-Patch Graph Convolutional Network for Image Denoising". in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. 2021.
- [73] Valsesia, D., G. Fracastoro, and E. Magli. "Image Denoising with Graph-Convolutional Neural Networks". in *2019 IEEE International Conference on Image Processing (ICIP)*. 2019.
- [74] Simonovsky, M. and N. Komodakis, *Dynamic Edge-Conditioned Filters in Convolutional Neural Networks on Graphs*. 2017. 29–38.
- [75] He, J., et al. "Conditional Sequential Modulation for Efficient Global Image Retouching". in *Computer Vision – ECCV 2020*. 2020. Cham: Springer International Publishing.
- [76] Chen, X., et al., *HDRUNet: Single Image HDR Reconstruction with Denoising and Dequantization*. 2021.
- [77] Tian, C., Y. Xu, and W. Zuo, "Image denoising using deep CNN with batch renormalization". *Neural Networks*, 2020. 121: p. 461–473.
- [78] Pan, J., et al. "Learning Dual Convolutional Neural Networks for Low-Level Vision". in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2018.
- [79] Helou, M.E. and S. Süsstrunk, "Blind Universal Bayesian Image Denoising With Gaussian Noise Level Learning". *IEEE Transactions on Image Processing*, 2020. 29: p. 4885–4897.
- [80] Kim, Y., J.W. Soh, and N.I. Cho, "Adaptively Tuning a Convolutional Neural Network by Gate Process for Image Denoising". *IEEE Access*, 2019. 7: p. 63447–63456.
- [81] He, J., C. Dong, and Y. Qiao. "Interactive Multi-dimension Modulation with Dynamic Controllable Residual Learning for Image Restoration". in *Computer Vision – ECCV 2020*. 2020. Cham: Springer International Publishing.
- [82] Li, B., et al. "All-In-One Image Restoration for Unknown Corruption". in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2022.
- [83] Xu, Y., et al., "Unified-Width Adaptive Dynamic Network for All-In-One Image Restoration". *ArXiv*, 2024. abs/2401.13221.
- [84] Soh, J.W. and N.I. Cho. "Deep Universal Blind Image Denoising". in *2020 25th International Conference on Pattern Recognition (ICPR)*. 2021.
- [85] Wang, W., et al. "CFSNet: Toward a Controllable Feature Space for Image Restoration". in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. 2019.
- [86] Holla, K.S. and B. Lee. "Convolutional Residual Blocks With Edge Guidance for Image Denoising". in *2022 13th International Conference on Information and Communication Technology Convergence (ICTC)*. 2022.
- [87] Tian, C., et al., "Designing and training of a dual CNN for image denoising". *Knowledge-Based Systems*, 2021. 226: p. 106949.
- [88] Yu, K., et al., "Path-Restore: Learning Network Path Selection for Image Restoration". *IEEE Transactions on*
- [58] Xie, J., L. Xu, and E. Chen, "Image Denoising and Inpainting with Deep Neural Networks". *Advances in Neural Information Processing Systems*, 2012. 1.
- [59] Wu, W., et al., *Image Blind Denoising Using Dual Convolutional Neural Network with Skip Connection*. 2023.
- [60] Mao, X.-J., C. Shen, and Y.-B. Yang, Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections, in *Proceedings of the 30th International Conference on Neural Information Processing Systems*. 2016, Curran Associates Inc.: Barcelona, Spain. p. 2810–2818.
- [61] Guo, S., et al. "Toward Convolutional Blind Denoising of Real Photographs". in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2019.
- [62] Ronneberger, O., P. Fischer, and T. Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation". in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. 2015. Cham: Springer International Publishing.
- [63] Liu, P., et al., "Multi-level Wavelet-CNN for Image Restoration". 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2018: p. 886–88609.
- [64] Zhang, K., et al., "Practical Blind Image Denoising via Swin-Conv-UNet and Data Synthesis". *Machine Intelligence Research*, 2023. 20(6): p. 822–836.
- [65] Fan, C.-M., T.-J. Liu, and K.-H. Liu, "SUNet: Swin Transformer UNet for Image Denoising". 2022 IEEE International Symposium on Circuits and Systems (ISCAS), 2022: p. 2333–2337.
- [66] Cao, H., et al. "Swin-Unet: Unet-Like Pure Transformer for Medical Image Segmentation". in *Computer Vision – ECCV 2022 Workshops*. 2023. Cham: Springer Nature Switzerland.
- [67] Lefkimmiatis, S. "Non-local Color Image Denoising with Convolutional Neural Networks". in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017.
- [68] Plötz, T. and S. Roth, Neural nearest neighbors networks, in *Proceedings of the 32nd International Conference on Neural Information Processing Systems*. 2018, Curran Associates Inc.: Montréal, Canada. p. 1095–1106.
- [69] Liu, D., et al., Non-local recurrent network for image restoration, in *Proceedings of the 32nd International Conference on Neural Information Processing Systems*. 2018, Curran Associates Inc.: Montréal, Canada. p. 1680–1689.
- [70] Qiao, P., et al., Learning Non-local Image Diffusion for Image Denoising, in *Proceedings of the 25th ACM international conference on Multimedia*. 2017, Association for Computing Machinery: Mountain View, California, USA. p. 1847–1855.
- [71] Valsesia, D., G. Fracastoro, and E. Magli, "Deep Graph-Convolutional Image Denoising". *IEEE Transactions on Image Processing*, 2020. 29: p. 8226–8237.

- [105] Liu, Z., et al. "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows". in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. 2021.
- [106] Liang, J., et al. "SwinIR: Image Restoration Using Swin Transformer". in *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*. 2021.
- [107] Dai, T., et al., Mix-order Attention Networks for Image Restoration, in *Proceedings of the 29th ACM International Conference on Multimedia*. 2021, Association for Computing Machinery: Virtual Event, China. p. 2880–2888.
- [108] Zhang, Y., et al., *Residual Non-local Attention Networks for Image Restoration*. 2019.
- [109] Wang, F., et al. "Residual Attention Network for Image Classification". in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017.
- [110] Wang, P., et al. "Understanding Convolution for Semantic Segmentation". in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*. 2018.
- [111] Du, Z., et al., "Multi-scale Attention Dilated Residual Image Denoising Network based on Skip Connection". *Instrumentation*, 2024. 11(3): p. 41–53.
- [112] Zhang, Y., et al., *KBNet: Kernel Basis Network for Image Restoration*. 2023.
- [113] Hassija, V., et al., "Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence". *Cognitive Computation*, 2024. 16(1): p. 45–74.
- [114] Hou, G., Y. Yang, and J.H. Xue. "Residual Dilated Network with Attention for Image Blind Denoising". in *2019 IEEE International Conference on Multimedia and Expo (ICME)*. 2019.
- [115] Tian, C., et al., "Enhanced CNN for image denoising". *CAAI Transactions on Intelligence Technology*, 2019. 4(1): p. 17–23.
- [116] Martin, D., et al. "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics". in *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*. 2001.
- [117] Deng, J., et al. "ImageNet: A large-scale hierarchical image database". in *2009 IEEE Conference on Computer Vision and Pattern Recognition*. 2009.
- [118] Ma, K., et al., "Waterloo Exploration Database: New Challenges for Image Quality Assessment Models". *IEEE Transactions on Image Processing*, 2017. 26(2): p. 1004–1016.
- [119] Franzen, R. Kodak lossless true color image suite. 1999; Available from: <https://r0k.us/graphics/kodak/>.
- Pattern Analysis and Machine Intelligence, 2022. 44(10): p. 7078–7092.
- [89] Yao, C., et al., "Dense Residual Transformer for Image Denoising". *Electronics*, 2022. 11: p. 418.
- [90] Vaswani, A., et al., Attention is all you need, in *Proceedings of the 31st International Conference on Neural Information Processing Systems*. 2017, Curran Associates Inc.: Long Beach, California, USA. p. 6000–6010.
- [91] Li, Y., et al. "Attention-Guided Unified Network for Panoptic Segmentation". in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2019.
- [92] Wu, W., et al., "Dual residual attention network for image denoising". *Pattern Recognition*, 2024. 149: p. 110291.
- [93] Woo, S., et al. "CBAM: Convolutional Block Attention Module". in *Computer Vision – ECCV 2018*. 2018. Cham: Springer International Publishing.
- [94] Hu, J., L. Shen, and G. Sun. "Squeeze-and-Excitation Networks". in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2018.
- [95] Wu, W., et al., "Dual convolutional neural network with attention for image blind denoising". *Multimedia Systems*, 2024. 30(5): p. 263.
- [96] Zhang, Y., et al., *Image Super-Resolution Using Very Deep Residual Channel Attention Networks: 15th European Conference, Munich, Germany, September 8–14, 2018, Proceedings, Part VII*. 2018. p. 294–310.
- [97] Anwar, S. and N. Barnes. "Real Image Denoising With Feature Attention". in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. 2019.
- [98] Islam, S., et al., "A comprehensive survey on applications of transformers for deep learning tasks". *Expert Systems with Applications*, 2024. 241: p. 122666.
- [99] Dosovitskiy, A., et al., "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale". *ArXiv*, 2020. abs/2010.11929.
- [100] Touvron, H., et al., Training data-efficient image transformers & distillation through attention, in *Proceedings of the 38th International Conference on Machine Learning*, M. Marina and Z. Tong, Editors. 2021, PMLR: Proceedings of Machine Learning Research. p. 10347--10357.
- [101] Carion, N., et al. "End-to-End Object Detection with Transformers". in *Computer Vision – ECCV 2020*. 2020. Cham: Springer International Publishing.
- [102] Radford, A., et al. "Learning Transferable Visual Models From Natural Language Supervision". in *International Conference on Machine Learning*. 2021.
- [103] Song, M., W. Wang, and Y. Zhao. "A Dynamic Network with Transformer for Image Denoising". *Electronics*, 2024. 13, DOI: 10.3390/electronics13091676.
- [104] Chen, Y., et al. "Dynamic Convolution: Attention Over Convolution Kernels". in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2020.



جواد آسلیمی ضامنجان مدرک کارشناسی و کارشناسی ارشد خود را در رشته مهندسی کامپیوتر از دانشگاه اراک دریافت کردند و در حال حاضر دانشجوی دکترای رشته مهندسی کامپیوتر در دانشگاه اراک می باشند. زمینه‌های پژوهشی مورد علاقه ایشان یادگیری ماشین و پردازش تصویر است.



محمد حسین شکور مدرک کارشناسی و کارشناسی ارشد خود را در رشته مهندسی کامپیوتر به ترتیب از دانشگاه شیراز و اصفهان دریافت کردند. ایشان مدرک دکترای خود را در رشته مهندسی کامپیوتر-گرایش هوش مصنوعی از دانشگاه شیراز اخذ نمودند. تخصص ایشان پردازش تصویر و یادگیری عمیق است و هم اکنون عضو هیات علمی دانشگاه اراک هستند.



محسن رحمانی مدرک کارشناسی خود را در رشته مهندسی کامپیوتر از دانشگاه شیراز اخذ کردند، ایشان مدارک کارشناسی ارشد و دکترای خود را در رشته مهندسی کامپیوتر از دانشگاه علم و صنعت ایران دریافت کردند. تخصص ایشان پردازش سیگنال و هوش مصنوعی است و هم اکنون عضو هیات علمی دانشگاه اراک هستند.