

الگوریتم‌های یادگیری عمیق در فراتفکیک‌پذیری تصاویر

بهار قادری^۱، حمید آزاد^{۲*}

۱- گروه مهندسی برق، واحد شیراز، دانشگاه آزاد اسلامی، شیراز، ایران

Bahar.ghaderi60@gmail.com

۲- گروه مهندسی برق، واحد شیراز، دانشگاه آزاد اسلامی، شیراز، ایران

azad@shirazu.ac.ir

چکیده: فراتفکیک‌پذیری تصویر یکی از فرآیندهای مهم پردازش تصویر جهت افزایش وضوح تصاویر و ویدئوها می‌باشد. در سال‌های اخیر، روش‌های مبتنی بر شبکه‌های عصبی عمیق جهت فراتفکیک‌پذیری شاهد پیشرفت قابل توجهی بوده است. هدف این مقاله ارائه یک بررسی جامع در مورد پیشرفت‌های اخیر فراتفکیک‌پذیری تصویر با استفاده از رویکردهای یادگیری عمیق است. در این مقاله، ضمن معرفی مفاهیم فراتفکیک‌پذیری تصویر، به بررسی الگوریتم‌های رایج یادگیری عمیق جهت فراتفکیک‌پذیری، و کاربردهای فراتفکیک‌پذیری پرداخته شده است. علاوه بر این، مجموعه پایگاه‌های داده و معیارهای ارزیابی تشریح داده می‌شود. این مقاله می‌تواند راهگشای محققان پردازش تصویر در فرآیند فراتفکیک‌پذیری باشد. اهتمام نویسندگان بر این بوده است که همه جنبه‌های این فرآیند مورد کاوش قرار گیرد.

واژه‌های کلیدی: فراتفکیک‌پذیری، الگوریتم‌های یادگیری عمیق، تصویر با کیفیت، شبکه عصبی کانولوشن.

Deep Learning Algorithms in Super-Resolution Images

Bahar Ghaderi¹, Hamid Azad^{2*}

¹ Department of Electrical Engineering, Shiraz Branch, Islamic Azad University, Shiraz, Iran

Bahar.ghaderi60@gmail.com

² Department of Electrical Engineering, Shiraz Branch, Islamic Azad University, Shiraz, Iran

azad@shirazu.ac.ir

Abstract:

Image super-resolution is one of the important image processing processes to increase the resolution of images and videos. In recent years, methods based on deep neural networks for super-resolution have seen significant progress. The aim of this paper is to provide a comprehensive review on recent developments in super-resolution image using deep learning approaches. In this article, while introducing the concepts of image super-resolution, the common deep learning algorithms for super-resolution and the applications of super-resolution have been investigated. In addition, the set of databases and evaluation criteria are described. This article can open the way for image processing researchers in the super-resolution process. The authors' effort has been to explore all aspects of this process.

Keywords: Super-resolution (SR), Deep Learning Algorithms, High Resolution (HR) Image, Low Resolution (LR) Image, Convolutional Neural Network.

DOI: 00.00000/0000

تاریخ چاپ مقاله: ۱۴۰۲/۰۲/۲۷

تاریخ پذیرش مقاله: ۱۴۰۲/۰۲/۱۳

نوع مقاله: پژوهشی

تاریخ ارسال مقاله: ۱۴۰۲/۰۱/۰۷

* نویسنده‌ی مسئول؛ نشانی: ایران - شیراز - دانشگاه آزاد اسلامی شیراز - دانشکده مهندسی - گروه مهندسی برق

۱- مقدمه

تقریبی \hat{I}_y از تصویر با کیفیت مرجع I_y و تصویر بی کیفیت I_x می‌باشد.

$$\hat{I}_y = F(I_x; \theta) \quad (2)$$

که F مدل فراتفکیک‌پذیری است و θ پارامترهای F می‌باشد. اگرچه فرآیند تخریب ناشناخته است و می‌تواند تحت تأثیر عوامل مختلفی مانند مصنوعات فشرده سازی، تخریب های ناهمسانگرد، نویز حسگر و نویز لکه‌ای قرار گیرد. محققان تلاش می‌کنند تا نقشه تخریب را مدل کنند. اکثراً مطابق (۳) به طور مستقیم تخریب را به عنوان یک عملیات نمونه برداری کوچک مدل می‌کنند.

$$D(I_y; \delta) = (I_y \otimes k) \downarrow_s + n_c \{k, s, \zeta\} \subset \delta \quad (3)$$

\downarrow_s عملگر نمونه‌برداری پایین با فاکتور مقیاس s می‌باشد. $(I_y \otimes k)$ کانولوشن بین کرنل محوشدگی و تصویر با کیفیت I_y و n_c مقداری نویز سفید گوسی با انحراف معیار ζ می‌باشد. بنابراین تابع هدف فراتفکیک‌پذیری مطابق (۴) است.

$$\hat{\theta} = \arg_{\theta} \min \mathcal{L}(\hat{I}_y, I_y) + \lambda \varphi(\theta) \quad (4)$$

در (۴)، $\mathcal{L}(\hat{I}_y, I_y)$ تابع اتلاف بین تصویر با کیفیت I_y و تصویر با کیفیت مرجع \hat{I}_y می‌باشد. $\varphi(\theta)$ ترم منظم‌سازی و λ پارامتر توازن است. توابع اتلاف مختلفی وجود دارد که محبوب‌ترین تابع اتلاف برای فرآیند فراتفکیک‌پذیری، خطای میانگین مربعات پیکسلی (اتلاف پیکسل) می‌باشد.

۳- الگوریتم‌های رایج یادگیری عمیق در فرآیند

فراتفکیک‌پذیری

الگوریتم‌های گوناگونی از شبکه‌های عصبی عمیق جهت فرآیند فراتفکیک‌پذیری تصاویر استفاده شده است که در این بخش، به شرح چند الگوریتم مهم و رایج پرداخته می‌شود.

۳-۱- شبکه عصبی کانولوشن^۳

شبکه‌های عصبی کانولوشن یکی از مهم‌ترین روش‌های یادگیری عمیق هستند که در آنها چندین لایه با روشی قدرتمند آموزش می‌بینند [۱۵]. این روش بسیار کارآمد بوده و یکی از رایج‌ترین روش‌ها در کاربردهای مختلف بینایی ماشین است. این الگوریتم از سه لایه اصلی شامل لایه کانولوشن، لایه ادغام و لایه تماماً متصل تشکیل شده است. لایه کانولوشن هسته اصلی شبکه می‌باشد که تصویر ورودی با کرنل های مختلف کانوالو می‌کند. لایه ادغام بعد از لایه کانولوشن قرار می‌گیرد که جهت کاهش اندازه نگاشت ویژگی‌ها، کاهش حجم محاسباتی و کنترل نویز استفاده می‌شود. لایه کانولوشن و لایه ادغام نسبت به انتقال تصاویر ثابت می‌باشند. لایه آخر این شبکه‌ها، لایه تماماً متصل می‌باشد. لایه ادغام و لایه تماماً متصل جهت فرآیند فراتفکیک‌پذیری

فراتفکیک‌پذیری تصویر، فرآیند بازیابی تصاویر وضوح بالا از تصاویر وضوح پایین می‌باشد [۱]. این فرآیند، مسئله مهم و چالش برانگیز در بینایی کامپیوتر و پردازش تصویر است [۲] و کاربرد گسترده‌ای در دنیای واقعی همچون بازسازی تصاویر پزشکی [۳]، سنجش از دور [۴]، سیستم‌های نظارتی [۵، ۶] و تلویزیون با کیفیت بالا [۷] دارد. روش های فراتفکیک‌پذیری به طور کلی به دو دسته، روش‌های سنتی و روش‌های یادگیری عمیق تقسیم می‌شود [۸]. روش‌های سنتی سال هاست که به طور گسترده استفاده می‌شود اما روش‌های یادگیری عمیق اخیراً استفاده شده‌اند که عملکرد بهتری دارند [۹-۱۲].

با توجه به اهمیت فرآیند فراتفکیک‌پذیری تاکنون مقاله مروری فارسی ارائه نشده است و همچنین تعداد معدودی مقالات انگلیسی وجود دارد که عمدتاً به بررسی روش‌های سنتی یا بررسی روش‌های یادگیری عمیق جهت فرآیند فراتفکیک‌پذیری بر روی نوع خاصی از تصاویر (پزشکی، چهره، سنجش از دور، طبیعی و غیره) پرداخته است [۴، ۵، ۷، ۸، ۱۳، ۱۴]. اهمیت و کاربرد روزافزون فرآیند فراتفکیک‌پذیری نشان می‌دهد که حداقل هر پنج سال یکبار باید یک مقاله مروری در این زمینه نوشته شود تا بتواند راهگشای محققان در جهت ارتقا عملکرد فرآیند فراتفکیک‌پذیری باشد. موارد ذکر شده، انگیزه بخش نگارش این مقاله در خصوص الگوریتم‌های یادگیری عمیق در فرآیند فراتفکیک‌پذیری بوده است. هدف این مقاله، تشریح مفهوم فرآیند فراتفکیک‌پذیری، بررسی الگوریتم‌های رایج یادگیری عمیق در فرآیند فراتفکیک‌پذیری، تشریح پایگاه تصاویر موجود و معیارهای ارزیابی می‌باشد.

سازمان‌دهی مقاله به این صورت است که در بخش دوم به بررسی فرآیند فراتفکیک‌پذیری و در بخش سوم به شرح الگوریتم‌های رایج یادگیری عمیق در فرآیند فراتفکیک‌پذیری پرداخته می‌شود. پایگاه‌های داده رایج، معیارهای ارزیابی، و کاربرد فراتفکیک‌پذیری به ترتیب در بخش چهارم، پنجم و ششم بررسی می‌شود. در بخش هفتم به جمع‌بندی و جهت‌دهی به آینده پرداخته می‌شود.

۲- فرآیند فراتفکیک‌پذیری

فرآیند فراتفکیک‌پذیری، تبدیل تصاویر بی کیفیت (کیفیت پایین) به تصاویر با کیفیت می‌باشد. فرآیند تخریب مطابق فرمول (۱) می‌باشد.

$$I_x = D(I_y; \delta) \quad (1)$$

در (۱)، D تابع تخریب و δ پارامترهای تخریب می‌باشد. در سناریوی دنیای واقعی، فقط I_x در دسترس است در حالی که هیچ اطلاعی درباره فرآیند تخریب یا پارامترهای تخریب وجود ندارد. هدف فراتفکیک‌پذیری، حذف اثر تخریب است و بازیابی تصاویر با کیفیت

³ Convolutional Neural Network (CNN)

¹High-Resolution (HR) Image

² Ground Truth HR Image



لازم نمی‌باشد، یعنی در این فرآیند تنها لایه کانولوشن کاربرد دارد [۱۶]. در هر شبکه عصبی کانولوشن دو مرحله آموزش وجود دارد، انتشار به جلو و پس‌انتشار که در مرحله انتشار به جلو تصویر ورودی وارد شبکه می‌شود. سپس عملیات کانولوشن انجام می‌شود و در نهایت خروجی شبکه محاسبه می‌شود. در این مرحله، نرخ خطای شبکه محاسبه می‌شود تا پارامترهای شبکه در مرحله بعدی تنظیم شوند. در مرحله پس‌انتشار، گرادیان هر پارامتر بر اساس قانون زنجیره محاسبه می‌شود و پارامترها با توجه به تأثیر روی خطای ایجاد شده در شبکه تغییر می‌کنند. در این شبکه، آموزش شبکه پس از انجام تعداد مناسب این مراحل به پایان می‌رسند.

در [۱۷]، شبکه عصبی کانولوشن جهت فراتفکیک‌پذیری تصاویر سی‌تی‌اسکن قفسه سینه معرفی شده است. در [۱۸]، روش شبکه‌های عصبی کانولوشن جهت فراتفکیک‌پذیری تصاویر ماهواره‌ای استفاده شده است. در [۱۹]، شبکه عصبی عمیق جهت فراتفکیک‌پذیری تصاویر کولونوسکوپی استفاده شده است. در [۲۰]، شبکه عصبی کانولوشنی، شبکه عمیق cascade و شبکه‌های کدبندی تُنک^۱ جهت فرآیند فراتفکیک‌پذیری تصاویر ماهواره‌ای بررسی شده است که شبکه عمیق cascade از نظر دقت هم‌ترازی و نسبت سیگنال به نویز عملکرد بهتری دارد. در [۲۱]، شبکه عصبی کانولوشن را جهت فراتفکیک‌پذیری تصویر بهبود داده است. در این روش جهت کاهش تعداد پارامترها از سه بلوک فرعی، بلوک استخراج و بهبود اطلاعات، بلوک بازسازی و بلوک پالایش اطلاعات در شبکه عصبی کانولوشن استفاده می‌کند. در [۲۲]، شبکه عصبی کانولوشن جهت افزایش سرعت در فرآیند فراتفکیک‌پذیری بهبود یافته است. شبکه باقیمانده و اتصال پرش در این شبکه استفاده می‌شود. در [۲۳]، شبکه عصبی کانولوشنی جهت فراتفکیک‌پذیری تصویر ارائه شده است که در این روش، نگاشت کلی بین تصاویر ورودی جهت ایجاد تصاویر با کیفیت استفاده می‌شود.

۲-۳- شبکه مولد متخاصم^۲

شبکه‌های مولد متخاصم، تکنیکی از روش‌های یادگیری ماشین محسوب می‌شوند. این الگوریتم در سال ۲۰۱۴ توسط یان گودفلو و همکاران مطرح و سپس در سال‌های اخیر مورد توجه پژوهشگران و فعال حوزه هوش مصنوعی قرار گرفت [۲۴]. در واقع شبکه‌های GAN رویکردی نوین برای مدل‌سازی مولد با بهره‌گیری از روش‌های یادگیری عمیق برپایه شبکه‌های عصبی کانولوشنی است. در ساختار الگوریتم‌های GAN مدل‌سازی مولد یک فعالیت نظارت نشده است، بدین معنی که یادگیری ماشین براساس اکتشاف خودکار و یادگیری قواعد یا الگوهای موجود در داده‌های ورودی می‌باشد. این تکنیک به

نحوی انجام می‌گردد، که مدل می‌تواند برای خلق یا تولید خروجی‌های جدید از داده‌های اصلی به نحوی عمل نماید که خروجی‌ها غیر اصلی شباهت بسیار بی‌نظیری به داده‌های اصلی داشته و در عین حال جدید و غیر تکراری باشند. شبکه‌های مولد تخصصی روشی بسیار هوشمندانه برای آموزش تولید مدل‌های جدید از داده‌های قبلی می‌باشند. در واقع الگوریتم در این روش بدین صورت کار می‌کند که با چهار چوب‌بندی مساله مورد نظر دو زیر مدل می‌سازد، مدل مولدی که وظیفه‌اش تولید نمونه‌های جدید است و دیگری مدل متمایزگر که تلاش می‌نماید تا نمونه‌ها را به عنوان نمونه اصلی و حقیقی یا نمونه جعلی دسته‌بندی نماید. این کار تا آن زمان ادامه پیدا می‌نماید که مدل متمایزگر به دفعات مختلف فریب خورده و دیگر نتواند داده‌های اصلی و غیراصلی را از هم جدا نماید.

در [۲۵]، روش جدیدی مبتنی بر شبکه‌های مولد متخاصم جهت فراتفکیک‌پذیری تصاویر سی‌تی‌اسکن کووید ۱۹ و تصاویر سی‌تی‌اسکن مغز پیشنهاد شده است. در [۲۶]، شبکه جدیدی به نام شبکه مولد متخاصم جهت فرآیند فراتفکیک‌پذیری تصاویر مغز ارائه شده است. در [۲۷]، شبکه متخاصم مولد چرخه^۳ جهت فرآیند فراتفکیک‌پذیری تصاویر سونوگرافی ارائه شده است که در این روش از چهار نوع تابع اتلاف شامل اتلاف پیکسل، اتلاف ویژگی‌های ادراکی^۴، اتلاف خصمانه^۵ و اتلاف ثبات چرخه^۶ استفاده شده است. در [۲۸]، شبکه متخاصم مولد تصادفی^۷ جهت فراتفکیک‌پذیری تصاویر همچون تصاویر شبکه‌ای و تصاویر قلب ارائه شده است. در [۲۹]، شبکه متخاصم مولد بهبودیافته^۸ جهت فرآیند فراتفکیک‌پذیری شبکه‌ای ارائه شده است. در این روش تابع اتلافی جدیدی که ترکیبی از توابع اتلاف نسبی خصمانه، اتلاف ادراکی و اتلاف میانگین مربعات خطا است، جهت آموزش شبکه پیشنهاد کرده است. در [۳۰]، شبکه تولید متمایز عمیق^۹ جهت فرآیند فراتفکیک‌پذیری تصاویر چهره ارائه شده است که این روش در تصاویر چهره کیفیت پایین (با تفکیک‌پذیری پایین) عملکرد مناسبی دارد.

۳-۳- شبکه عصبی بازگشتی

یکی از الگوریتم‌های یادگیری عمیق که دارای حافظه درونی هستند، شبکه عصبی بازگشتی است [۳۱]. پیش‌فرض سایر الگوریتم‌های یادگیری عمیق اینست که بین داده‌های ورودی شبکه، هیچ گونه وابستگی وجود ندارد و داده‌ها از هم مستقل هستند. شبکه عصبی بازگشتی با داشتن یک حافظه درونی، مناسب مسائلی هستند که داده‌های ورودی آن‌ها به یکدیگر وابسته است و باید به ترتیب ورودی این داده‌ها توجه شود. مدل شبکه عصبی بازگشتی از حافظه درونی خود برای نگهداری داده‌های ورودی قبلی استفاده می‌کند تا بر اساس آن‌ها

⁶ Cycle Consistency Loss

⁷ Progressive Generative Adversarial Networks

⁸ Improved Generative Adversarial Networks

⁹ Deep discriminative Generative Network

¹ Sparse Coding Network

² Generative Adversarial Networks (GAN)

³ Cycle Generative Adversarial Network(CycleGAN)

⁴ Perceptual Feature Loss

⁵ Adversarial Loss



های دیگر شبکه عصبی عمیق، پارامترهای کمتری دارد و عملکرد موفق-تری جهت فرایند فراتفکیک‌پذیری دارد. در [۴۲]، شبکه عصبی کانولوشن-حافظه طولانی کوتاه مدت^۶ جهت فراتفکیک‌پذیری تصاویر ارائه شده است. در [۴۳]، شبکه‌های عصبی حافظه طولانی کوتاه مدت کانولوشن جهت فراتفکیک‌پذیری تصویر پیشنهاد شده است. در [۴۱]، شبکه LSTM کانولوشن دو طرفه^۸ جهت فراتفکیک‌پذیری تصویر سنجش از دور و ماهواره‌ای ارائه شده است. این روش از پارامترهای کمتری استفاده می‌کند اما عملکرد مناسب‌تری نسبت به روش‌های کلاسیک دارد. در [۴۴]، روشی مبتنی بر شبکه حافظه طولانی کوتاه مدت جهت فراتفکیک‌پذیری تصاویر فراطیفی پیشنهاد شده است. این روش، ویژگی‌های فضایی و ویژگی‌های طیفی تصاویر فراطیفی را در نظر می‌گیرد که از جهت دقت فرایند فراتفکیک‌پذیری عملکرد بهتری دارد.

۳-۵- شبکه باقیمانده

شبکه باقیمانده توسط شیوکینگ رن و همکاران در سال ۲۰۱۵ ارائه شده است. شبکه عصبی باقیمانده جهت حل مشکل انفجار گرادیان از لایه‌های پرشی و میانبر بین لایه‌ها استفاده می‌کند [۴۵]. این شبکه در ساختار پایه شامل ۳۴ لایه ساده با معماری VGGNet است که به آن اتصالات میان‌بر اضافه شده است. شبکه عصبی باقیمانده در نسخه-های ۵۰، ۱۰۱ و ۱۵۲ لایه هم وجود دارد. در [۴۶]، شبکه عصبی باقیمانده چند مقیاسی^۹ و تابع هدف خطای میانگین مربعات^{۱۰} جهت فراتفکیک‌پذیری تصاویر ماهواره‌ای پیشنهاد شده است. این روش سبب می‌شود، اطلاعات فرکانس بالا در تصاویر همچون لبه‌ها و بافت‌ها حفظ شوند. در [۴۷]، شبکه عصبی باقیمانده جهت فراتفکیک‌پذیری تصاویر MRI ارائه شده است. در این روش، درون‌یابی دو مکعبی با یک لایه دکانولوشن جهت یادگیری فیلترهای نمونه‌برداری بالا استفاده می‌شود و از لایه‌های اتصال پرش باقیمانده با بلوک Inception جهت باز یابی تصویر با کیفیت از تصویر بی کیفیت به کار گرفته می‌شود. در [۴۸]، روش مبتنی بر شبکه عصبی باقیمانده جهت فرایند فراتفکیک‌پذیری پیشنهاد شده است. در این روش، ابتدا تصویر ورودی توسط تبدیل موجک گسسته به چهار زیر تصویر تجزیه می‌شود و سپس ضرایب چهار زیر تصویر به شبکه عصبی باقیمانده داده می‌شوند و در نهایت از معکوس تبدیل موجک گسسته جهت باز یابی تصویر با کیفیت استفاده می‌شود. در [۴۹]، رویکرد جدیدی مبتنی بر شبکه باقیمانده جهت فراتفکیک‌پذیری تصاویر ارائه شده است. شبکه متراکم با توجه کانال باقیمانده عمیق^{۱۱} از اتصالات متراکم بین گروه‌های باقیمانده استفاده می‌کند که این روش عملکرد خوبی دارد.

به یادگیری داده جدید بپردازد. شبکه‌های بازگشتی شامل شبکه کانولوشن بازگشتی عمیق، شبکه باقیمانده بازگشتی عمیق و شبکه فراتفکیک‌پذیری هرم لاپلاسی عمیق^۱ به طور گسترده در فرایند فراتفکیک‌پذیری استفاده شده است.

در [۳۲]، مدل DRRCAN^۲ جهت فراتفکیک‌پذیری تصویر پیشنهاد شده است. در این روش با استفاده از بلوک بازگشتی سبب کاهش تعداد پارامترها و با استفاده از ماژول باقیمانده و اتصالات پرش سبب جلوگیری انفجار گرادیان، بهبود پایداری و توانایی تعمیم مدل می‌شود. در [۳۳]، شبکه عصبی بازگشتی جهت فراتفکیک‌پذیری استفاده شده است که این روش در برابر پایگاه داده واقعی و مصنوعی عملکرد موفق دارد. در [۳۴]، شبکه عصبی بازگشتی دو حالتی^۳ جهت فراتفکیک‌پذیری تصویر ارائه شده است. این روش نسبت به روش شبکه عصبی بازگشتی تک حالتی^۴ از جهت دقت عملکرد بهتری دارد. در [۳۵]، شبکه سه بعدی شبه بازگشتی دو طرفه^۵ برای فراتفکیک‌پذیری تصاویر فراطیفی ارائه شده است. این رویکرد دقت باز یابی و کیفیت بصری موثری دارد. در [۳۶]، شبکه عصبی بازگشتی جهت فراتفکیک‌پذیری تصاویر سنجش از دور استفاده شده است و این روش نسبت به روش‌های کلاسیک عملکرد بهتری دارد. در [۳۷]، شبکه بازگشتی عمیق جهت فرایند فراتفکیک‌پذیری پیشنهاد شده است. این روش جهت کاهش دشواری‌های آموزش از نظارت بازگشتی و اتصال پرش استفاده می‌کند. در [۳۸]، شبکه فراتفکیک‌پذیری هرم لاپلاسی عمیق ارائه شده است. این روش جزء شبکه عصبی بازگشتی می‌باشد که از نظر زمان اجرا و کیفیت فرایند فراتفکیک‌پذیری عملکرد مناسبی دارد.

۳-۴- شبکه حافظه طولانی کوتاه مدت^۶

شبکه‌های حافظه طولانی کوتاه مدت توسط هوکرایتر و اشمیدوبر در سال ۱۹۹۷ اختراع شدند و رکوردهای دقت را در حوزه‌های کاربردی متعدد ثبت کردند [۳۹، ۴۰]. شبکه حافظه طولانی کوتاه مدت از مشکل انفجار گرادیان جلوگیری می‌کند. شبکه عصبی با حافظه طولانی کوتاه مدت ساختار حلقه‌ای دارد و به ازای تعداد اجزای داده ورودی، عمل یادگیری و پردازش داده‌ها تکرار می‌شود. شبکه عصبی با حافظه طولانی کوتاه مدت داده‌های غیر مرتبط و بی‌اهمیت قبلی را در زمان پردازش داده جدید حذف می‌کند و تنها اطلاعات مهم داده‌های قبلی را در حافظه خود نگه می‌دارد. چنانچه نیاز به بروز رسانی اطلاعات قبلی نیز باشد، این کار توسط توابع درون این شبکه انجام می‌شود.

در [۴۱]، از شبکه حافظه طولانی کوتاه مدت جهت فراتفکیک‌پذیری تصاویر سنجش از دور استفاده شده است. این روش نسبت به الگوریتم-

⁶ Long short-term memory (LSTM)

⁷ CNN-LSTM network

⁸ Bidirectional convolutional LSTM (BiConvLSTM)

⁹ Multi-Scale Residual Neural Network (MRNN)

¹⁰ Mean Squared Error

¹¹ DenseNet with deep Residual Channel Attention (DRCA)

¹ Deep Laplacian Pyramid Super-Resolution Network

² Deep Recursive Residual Channel Attention Network (DRRCAN) model

³ Dual-State Recurrent Network (DSRN)

⁴ Single-state recurrent neural network (RNN)

⁵ Bidirectional 3D Quasi-Recurrent Neural Network



۴- پایگاه‌های داده

مجموعه به ترتیب شامل ۸۸۴، ۷۸۳ و ۸۴۰ جفت تصاویر باکیفیت و بی‌کیفیت در فاکتورهای مقیاس مختلف ۲، ۳ و ۴ می‌باشد.

• مجموعه City 100

مجموعه City 100 توسط چن و همکاران با دوربین‌های iPhoneX و NikonD5500 تهیه شده است [۵۴]. این مجموعه شامل ۱۰۰ تصویر از صحنه‌های شهری مختلف می‌باشد. تصاویر با کیفیت و بی‌کیفیت با فاصله کانونی متفاوتی گرفته شده‌اند.

• مجموعه SR-RAW

مجموعه SR-RAW توسط ژانک و همکاران ایجاد شده است که از هر صحنه، هفت تصویر تحت شرایط زوم مختلف (۲۴-۲۴۰) تهیه شده است [۵۵]. این مجموعه شامل ۵۰۰ توالی هفت تصویر از صحنه‌های داخلی و بیرونی می‌باشد.

• مجموعه TextZoom

مجموعه تصاویر متنی TextZoom از مجموعه RealSR و SR-RAW نشات گرفته است [۵۶]. تصاویر متنی در این مجموعه از تصاویر صحنه‌های مختلف RealSR و SR-RAW همچون مغازه‌ها، نماهای خیابان، و فضای داخلی خودرو برش داده شده‌اند. تصاویر متنی براساس دشواری به سه سطح، آسان، متوسط و سخت می‌باشد.

• مجموعه Super

مجموعه Super توسط کوهر و همکاران ایجاد شده است [۵۷]. این مجموعه بیش از ۸۰۰۰ تصویر از ۱۴ صحنه آزمایشگاهی در چهار وضوح و پنج سطح فشرده سازی با دوربین CMOS گرفته شده است.

• مجموعه ImagePairs

مجموعه ImagePairs توسط جوز و همکاران تهیه شده است که این مجموعه شامل ۱۱۴۲۱ جفت تصاویر با کیفیت و بی‌کیفیت از صحنه‌های مختلف گرفته شده است [۵۸]. این تصاویر با دو دوربین جهت تولید تصاویر با کیفیت و بی‌کیفیت گرفته شده است که فاصله کانونی متفاوتی دارند.

۴-۲- پایگاه تصاویر پزشکی

پایگاه‌های داده گوناگونی در مقالات مختلف جهت فرآیند فراتفکیک پذیری تصاویر پزشکی استفاده شده‌اند که تصاویر کیفیت پایین در بعضی از این پایگاه‌های داده به صورت تجربی و در بعضی دیگر شبیه‌سازی شده است. در ادامه چهار پایگاه تصاویر پزشکی جهت فرآیند فراتفکیک‌پذیری به طور مختصر توضیح داده می‌شود.

- پایگاه تصاویر MRI مغز شبیه‌سازی شده که شامل تصویر MRI مغز سالم و تصاویر MRI مغز بیماران مالتیپل اسکلروزیس می‌باشد. این تصاویر MRI در سه مودالیت، T_1 ، T_2 و PD می‌باشد^۴.

فرآیند فراتفکیک‌پذیری، یک شاخه علمی مهم در پردازش تصویر می‌باشد و برای ارزیابی عملکرد این فرآیند، تعدادی پایگاه داده استاندارد ایجاد شده است. از نظر ماهیت تصاویر، می‌توان این پایگاه‌ها را به چهار دسته، پایگاه تصاویر طبیعی، پزشکی، ماهواره‌ای، و پایگاه تصاویر چهره قابل تقسیم کرده است. در ادامه به بررسی هر یک از انواع پایگاه‌های داده پرداخته می‌شود.

۴-۱- پایگاه تصاویر طبیعی

پایگاه داده مختلفی از تصاویر طبیعی وجود دارد که بعضی از این پایگاه‌های داده شامل تصاویر با کیفیت و بی‌کیفیت هستند و بعضی دیگر از آنها تنها تصاویر با کیفیت تامین می‌کنند که تصاویر بی‌کیفیت در آنها می‌توان توسط نرم‌افزارهای مختلف ایجاد کرد. در ادامه به بررسی چند پایگاه داده رایج تصاویر طبیعی پرداخته می‌شود.

• مجموعه DIV2KRRK

این مجموعه داده آزمایش مصنوعی توسط بیل کیلگر و همکاران ساخته شده است [۵۰]. DIV2KRRK از پایگاه داده DIV2K نشات گرفته است. DIV2K توسط آگوستسون و همکاران ایجاد شد که ۱۰۰ عدد از تصاویر با کیفیت برای ایجاد تصاویر بی‌کیفیت (کیفیت پایین) با کرنل‌های تصادفی محو و نمونه‌برداری می‌شوند [۵۱]. کرنل‌های محوشدگی تصادفی با اندازه 11×11 گوسی ناهمسانگرد می‌باشد و هر کدام از آنها با دو طول $0/6$ و 5 توزیع شده مستقل و زاویه چرخش π و $-\pi$ می‌باشد. مدل تخریب در مجموعه DIV2KRRK، پیچیده‌تر و تصادفی‌تر از مجموعه DIV2K می‌باشد.

• مجموعه RealSR

مجموعه داده واقعی RealSR توسط چیا و همکاران جمع‌آوری شده است [۵۲]. داده آموزشی و آزمایشی مجموعه شامل ۵۹۵ جفت تصویر باکیفیت و بی‌کیفیت می‌باشد. این تصاویر با فاکتورهای مقیاس مختلف (۲، ۳، ۴) و در چهار طول فاصله کانونی (۲۸ میلی‌متر، ۳۵ میلی‌متر، ۵۰ میلی‌متر و ۱۰۵ میلی‌متر) گرفته شده‌اند. تصاویر در فاصله کانونی ۱۰۵ میلی‌متر به عنوان تصاویر با کیفیت و تصاویر در فواصل کانونی ۲۸ میلی‌متر، ۳۵ میلی‌متر و ۵۰ میلی‌متر به عنوان تصاویر بی‌کیفیت در نظر گرفته شده است.

• مجموعه DRealSR

مجموعه داده واقعی DRealSR توسط ویی و همکاران گرفته شده است [۵۳]. این مجموعه با پنج دوربین (سونی، کانن^۱، الپوس^۲، نیکون^۳ و پاناسونیک) در چهار وضوح از صحنه‌های بیرونی و داخلی (ساختمان‌ها، ادارات، گیاهان، پوسترها و غیره) تهیه شده است. این

³ Nikon

⁴ <http://brainweb.bic.mni.mcgill.ca/brainweb/>

¹ Canon

² Olympus



معیارهای ارزیابی مختلف ممکن است لزوماً با یکدیگر و همچنین ارزیابی ذهنی سازگار نباشند [۵۲]. در ادامه، معیارهای ارزیابی رایج عینی مبتنی بر کیفیت شرح داده می‌شود.

• نسبت سیگنال به نویز

معیار نسبت سیگنال به نویز مطابق (۵) محاسبه می‌شود. هرچه مقدار این معیار بیشتر، عملکرد فرآیند فراتفکیک پذیری بهتر می‌باشد.

$$PSNR = 10 \log_{10} \left(\frac{M^2}{MSE} \right) \quad (5)$$

در این رابطه، M ماکزیمم مقدار شدت روشنایی پیکسل (برای مثال برای تصاویر ۸ بیتی، مقدار ۲۵۵) می‌باشد. MSE میانگین مربعات خطا بین تصویر حاصل فرآیند فراتفکیک‌پذیری و تصویر با کیفیت می‌باشد که (۶) محاسبه می‌شود.

$$MSE = \frac{1}{hwc} \|I_{SR} - I_{HR}\|^2 \quad (6)$$

در این روابط، h, w, c به ترتیب کانال تصاویر، پهنا و ارتفاع می‌باشد. هرچه تفاضل پیکسل‌های بین دو تصویر (تصویر حاصل فرآیند فراتفکیک‌پذیری و تصویر باکیفیت) کمتر باشد، مقدار نسبت سیگنال به نویز بیشتر و بهتر می‌باشد.

• شاخص شباهت ساختاری

معیار شاخص شباهت ساختاری مطابق (۷) محاسبه می‌شود [۵۹].

$$SSIM = l(I_{HR}, I_{SR}) * C(I_{HR}, I_{SR}) * S(I_{HR}, I_{SR}) \quad (7)$$

در این رابطه، $l(I_{HR}, I_{SR})$ ، $C(I_{HR}, I_{SR})$ و $S(I_{HR}, I_{SR})$ به ترتیب شباهت شدت روشنایی، شباهت کنتراست و شباهت ساختاری می‌باشد. هرچه شباهت ساختاری دو تصویر بیشتر باشد، مقدار شاخص شباهت ساختاری بیشتر و بهتر می‌باشد. در [۵۹]، معیارهای ارزیابی نسبت سیگنال به نویز و شاخص شباهت ساختاری بررسی شدند که شاخص شباهت ساختاری کیفیت بصری را بهتر از معیار نسبت سیگنال به نویز منعکس می‌کند.

• معیار وفاداری اطلاعات^{۱۱}

معیار وفاداری اطلاعات، جزء معیارهای مرجع کامل است که کیفیت تصاویر را بر اساس آمار صحنه‌های طبیعی ارزیابی می‌کند. به طور کلی، تحریف‌ها، آمار صحنه‌های طبیعی را مختل می‌کند و تصاویر را غیرطبیعی نشان می‌دهد. معیار وفاداری اطلاعات، اطلاعات متقابل شرطی بین تصویر منبع و تصویر منحرف شده را مطابق (۸) محاسبه می‌کند [۶۰].

$$IFC = \sum_{k \in \text{subbands}} I(C^{N_k, k}, D^{N_k, k} | S^{N_k, k}) \quad (8)$$

- پایگاه تصاویر MRI مغز می‌باشد که این پایگاه داده در سه دسته، تصاویر MRI مغز افراد سالم، تصاویر MRI مغز بیماران آلزیمری و تصاویر MRI مغز نوزدان نارس می‌باشد. این پایگاه شامل ۳۱۲ تصاویر می‌باشد.^۱
- پایگاه تصاویر MRI مغز بیماران گلیوما می‌باشد و شامل ۶۵ تصویر می‌باشد.^۲
- پایگاه تصاویر MRI مغز افراد سالم می‌باشد که این پایگاه از سه بیمارستان مختلف با دستگاه مختلف MRI در لندن جمع‌آوری شده است و تعداد کل این تصاویر تقریباً برابر ۶۰۰ تصویر می‌باشد. این تصاویر در مودالیت‌های T_1 ، T_2 و PD می‌باشد.^۳

۳-۴- پایگاه تصاویر چهره

پایگاه‌های داده گوناگونی جهت فرآیند فراتفکیک‌پذیری تصاویر چهره وجود دارد که پایگاه Helen و پایگاه CelebA-HQ دو نمونه رایج می‌باشد. پایگاه‌های دیگری، جهت فرآیند فراتفکیک‌پذیری استفاده می‌شود.^۴

۴-۴- پایگاه تصاویر سنجش از دور

پایگاه‌های مختلفی جهت فرآیند فراتفکیک‌پذیری تصاویر ماهواره‌ای وجود دارد که در ادامه به شرح پایگاه‌های داده رایج پرداخته می‌شود.

- پایگاه RSSCN7: این پایگاه شامل ۲۸۰۰ تصویر از مناطق مسکونی، صنعتی، جنگل‌ها، زمین‌های کشاورزی، رودخانه‌ها و دریاچه‌ها می‌باشد.^۵
- پایگاه PIRM: این پایگاه شامل ۲۰۰ تصویر از محیط، مردم، گیاهان، مناظر طبیعی و اشیاء می‌باشد.^۶
- پایگاه WHURS19: این پایگاه شامل ۱۰۰۵ تصاویر از فرودگاه، پل، جنگل‌ها، مناطق تجاری، مسکونی و صنعتی می‌باشد.^۷
- پایگاه Kaggle: این پایگاه شامل ۳۲۴ تصویر از ساختمان‌ها، زمین گلف، جنگل‌ها و بندرگاه می‌باشد.^۸

۵- معیارهای ارزیابی

معیارهای ارزیابی کیفیت فرآیند فراتفکیک‌پذیری را می‌توان به طور کلی به دو دسته، ارزیابی ذهنی مبتنی بر ادراک انسانی^۹ و ارزیابی عینی مبتنی بر کیفیت^{۱۰} تقسیم کرد. محدودیت‌های ارزیابی ذهنی عبارتند از، نتیجه ارزیابی تحت تأثیر ترجیحات شخصی قرار می‌گیرد، فرآیند ارزیابی اغلب پرهزینه است و نمی‌توان آن را خودکار کرد. در مقابل، ارزیابی عینی برای استفاده راحت‌تر است، اگرچه نتایج با

⁷ http://www.tsi.enst.fr/~xia/satellite_image_project.html

⁸ <https://www.kaggle.com/c/drapper-satellite-image-chronology/data>

⁹ human perception-based subjective evaluation

¹⁰ quality metrics-based objective evaluation

¹¹ Information Fidelity Criterion

¹ <https://sites.google.com/site/brainseg/>

² <https://www.smir.ch/BRATS/Start2015>

³ <http://brain-development.org/ixi-dataset/>

⁴ <http://vis-www.cs.umass.edu/lfw/>

⁵ <https://sites.google.com/site/qinzoucn/documents>

⁶ <https://pirm.github.io/>



در این رابطه، $C^{N_k, k}$ برابر با N_k تا ضرایب از مدل تصویر منبع از k^{th} زیر باند می‌باشد و $D^{N_k, k}$ برابر N_k تا ضرایب از مدل تصویر منحرف شده از k^{th} زیر باند می‌باشد.

• معیار شباهت وصله تصویر ادراکی آموخته شده^۱

شباهت وصله تصویر ادراکی آموخته شده جهت ارزیابی کیفیت تصویر مبتنی بر مرجع است [۶۱]. این معیار فواصل بین تصویر مرجع و تصویر آزمایشی در یک فضای ویژگی عمیق محاسبه می‌کند. این معیار بیشتر از معیارهای ارزیابی نسبت سیگنال به نویز و شاخص شباهت ساختاری تطابق با قضاوت‌های انسانی (ارزیابی بصری) دارد. هرچه دو تصویر شبیه‌تر باشند، مقدار این معیار کمتر می‌باشد.

• ارزیاب کیفیت تصاویر طبیعی^۲

ارزیاب کیفیت تصویر طبیعی، یک معیار کاملاً کور بدون آگاهی از قضاوت یا تعریفات انسانی است [۶۲]. به طور خاص، ویژگی‌ها شامل پارامترهای توزیع گوسی تعمیم یافته^۳ و توزیع گوسی تعمیم یافته نامتقارن^۴ است که رفتار تکه‌های تصویر را مشخص می‌کند. سپس، کیفیت یک تصویر با استفاده از فاصله بین دو مدل گوسی چند متغیره متناسب با تصاویر طبیعی و تصویر ارزیابی شده اندازه‌گیری می‌شود.

• معیار کیفیت بدن مرجع^۵

این معیار جهت ارزیابی تصویر فراتفکیک‌پذیری شده استفاده می‌شود. برای پیش‌بینی مقدار ادراکی تصاویر فراتفکیک‌پذیری شده، سه گروه از ویژگی‌های آماری شامل ویژگی‌های فرکانس محلی، ویژگی‌های فرکانس عمومی و ویژگی‌های فضایی استخراج می‌شوند [۶۳].

• ارزیاب کیفیت مبتنی بر ادراک^۶

ارزیاب کیفیت مبتنی بر ادراک، یک معیار ارزیابی کیفیت تصویر بدون مرجع است [۶۴]. با توجه به اینکه توجه سیستم بینایی انسان به شدت به سمت مناطق فعال فضایی معطوف شده است، تصویر آزمایشی به بلوک‌های غیرهمپوشانی تقسیم می‌شود و تجزیه و تحلیل سطح بلوک برای شناسایی اعوجاج و کیفیت درجه انجام می‌شود. بنابراین، ارزیاب کیفیت مبتنی بر ادراک می‌تواند یک نقشه کیفیت فضایی ارائه دهد. کیفیت کلی تصویر ارزیابی شده را می‌توان با ادغام نمرات کیفیت سطح بلوک به دست آورد.

۶- کاربرد فراتفکیک‌پذیری

فرآیند فراتفکیک‌پذیری در زندگی روزمره کاربرد گسترده‌ای دارد که در سه دهه اخیر، کاربردهای مختلف آن مورد توجه محققان قرار گرفته است. در ادامه به شرح چند کاربرد مهم این حوزه پرداخته می‌شود.

۶-۱- ارتقا اطلاعات ویدیویی

کاربرد تکنیک‌های فراتفکیک‌پذیری به زندگی روزمره وارد شده است. تصاویر ویدیویی بی‌کیفیت را می‌توان با استفاده از این فرآیند به تصاویر ویدیویی با کیفیت تبدیل کرد. هیتاچی با استفاده از فراتفکیک‌پذیری، تبدیل تلویزیون با کیفیت استاندارد^۷ به تلویزیون با کیفیت بالا^۸ را به دست آورد. همین موضوع سبب شد که فراتفکیک‌پذیری را به موضوع تحقیقاتی بسیار جالب و مرتبط تبدیل کند [۶۵].

۶-۲- نظارت

امروزه دستگاه‌های ضبط ویدئوی دیجیتال^۹ در همه جا وجود دارند و نقش مهمی در کاربردهایی همچون نظارت بر ترافیک و نظارت بر امنیت ایفا می‌کنند. با این حال، در حال حاضر تجهیز دستگاه‌های منابع انسانی در مقیاس بزرگ غیرممکن است. بنابراین، مطالعه تکنیک‌های تصویر ضروری است. اگرچه تکنیک‌ها به تدریج توسعه یافته‌اند، استفاده عملی از ویدئو فراتفکیک‌پذیری هنوز یک چالش است. اولاً، دستگاه‌های ویدیویی در فضای باز در برابر تأثیر شرایط آب و هوایی آسیب‌پذیر هستند. علاوه بر این، داده‌های ویدیویی معمولاً دارای حجم عظیمی از داده و حرکت پیچیده هستند [۶۶].

۶-۳- تشخیص پزشکی

مودالیت‌های مختلف تصویربرداری پزشکی می‌توانند هم اطلاعات تشریحی در مورد ساختار بدن انسان و هم اطلاعات عملکردی را ارائه دهند. با این حال، محدودیت وضوح همیشه ارزش تصاویر پزشکی در تشخیص را کاهش می‌دهد. فناوری فراتفکیک‌پذیری با مودالیت‌های مهم تصویربرداری پزشکی همچون MRI، fMRI و PET استفاده شده- اند [۶۷].

۶-۴- رصد نجومی

وضوح فیزیکی دستگاه‌های تصویربرداری نجومی محدود شده توسط پارامترهای سیستم نیز فرصتی را برای تکنیک‌های فراتفکیک‌پذیری برای ایفای نقش فراهم می‌کند [۶۸]. سیستم‌های نجومی معمولاً می‌توانند مجموعه‌ای از تصاویر را برای فراتفکیک‌پذیری جمع‌آوری کنند. با بهبود وضوح تصاویر نجومی، فراتفکیک‌پذیری می‌تواند به اخترشناسان در اکتشاف فضای بیرونی کمک کند.

⁶ Perception-based Quality Evaluator (PIQE)

⁷ Standard Definition TV (SDTV)

⁸ High-Definition Television (HDTV)

⁹ Digital Video Recorder (DVR) Devices

¹ The Learned Perceptual Image Patch Similarity (LPIPS)

² Natural Image Quality Evaluator (NIQE)

³ Generalized Gaussian Distribution (GGD)

⁴ Asymmetric Generalized Gaussian distribution (AGGD)

⁵ Learned no-Reference Quality Metric (NRQM)

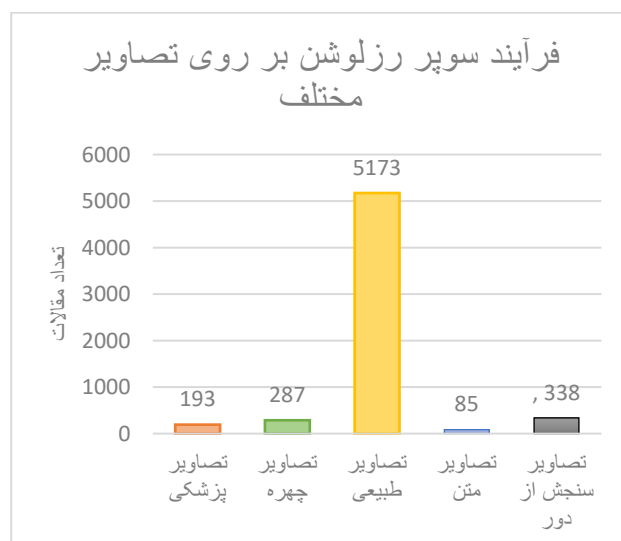


۶-۵- شناسایی اطلاعات بیومتریک

فرآیند فراتفکیک‌پذیری در تشخیص بیومتریک همچون ارتقای وضوح چهره [۶۹]، اثر انگشت [۷۰] و عنیبه [۷۱] مهم است. وضوح تصاویر بیومتریک در فرآیند شناسایی و تشخیص بسیار مهم است.

۷- جمع‌بندی و جهت‌دهی آینده

فراتفکیک‌پذیری تصویر، رویکردی موثر جهت افزایش وضوح فضایی برای کاربردهای متعددی همچون تشخیص اشیاء و طبقه‌بندی می‌باشد. این مقاله به طور خلاصه، الگوریتم‌های رایج یادگیری عمیق همچون شبکه عصبی کانولوشن، شبکه مولد متخاصم، شبکه بازگشتی و شبکه باقیمانده جهت فراتفکیک‌پذیری تصاویر بررسی می‌کند. علاوه بر این، به مجموعه پایگاه‌های داده رایج تصاویر مختلف و معیارهای ارزیابی پرداخته شد. مطالب ارائه شده در این مقاله، پتانسیل بررسی جهت سوگیری تحقیقات در زمینه فراتفکیک‌پذیری را دارد. با توجه به تحقیقات زیادی که در زمینه فراتفکیک‌پذیری انجام شده است، اما هنوز به دلیل متنوع بودن تصاویر و وجود مدل‌های تخریب مختلف در تصاویر، یک روش عمومی و کامل جهت انجام این فرآیند وجود ندارد. بنابراین برای انجام فراتفکیک‌پذیری تصاویر باید به ماهیت تصاویر هم توجه کرد، زیرا مدل‌های تخریب متفاوت بستگی به نوع تصویر دارد. شکل (۱) نمودار تعداد مقالات از پایگاه IEEE در تصاویر متفاوت در محدوده‌ی سال‌های ۲۰۱۵-۲۰۲۳ است. با توجه به این نمودار می‌توان نتیجه گرفت که در آینده زمینه‌های کاری بر روی تصاویر طبیعی و سنجش از دور بیشتر است. همانطور که مشخص شد، کاربردهای فراتفکیک‌پذیری در تصاویر طبیعی و تصاویر سنجش از دور بیشتر از تصاویر دیگر همچون تصاویر متنی و پزشکی مورد اقبال محققین است.



شکل (۱): نمودار تعداد مقالات بر حسب تصاویر مختلف فراتفکیک‌پذیری

مراجع

- [1] F. Liu, X. Yang, and B. De Baets, "A deep recursive multi-scale feature fusion network for image super-resolution," *Journal of Visual Communication and Image Representation*, vol. 90, p. 103730, 2023.
- [2] L. Inzerillo, F. Acuto, G. Di Mino, and M. Z. Uddin, "Super-resolution images methodology applied to UAV datasets to road pavement monitoring," *Drones*, vol. 6, p. 171, 2022.
- [3] Y. Huang, L. Shao, and A. F. Frangi, "Simultaneous super-resolution and cross-modality synthesis of 3D medical images using weakly-supervised joint convolutional sparse coding," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 6070-6079.
- [4] D. Yang, Z. Li, Y. Xia, and Z. Chen, "Remote sensing image super-resolution: Challenges and approaches," in *2015 IEEE international conference on digital signal processing (DSP)*, 2015, pp. 196-200.
- [5] J. Jiang, C. Wang, X. Liu, and J. Ma, "Deep learning-based face super-resolution: A survey," *ACM Computing Surveys (CSUR)*, vol. 55, pp. 1-36, 2021.
- [6] A. B. Deshmukh and N. Usha Rani, "Fractional-Grey Wolf optimizer-based kernel weighted regression model for multi-view face video super resolution," *International Journal of Machine Learning and Cybernetics*, vol. 10, pp. 859-877, 2019.
- [7] H. Liu, Z. Ruan, P. Zhao, C. Dong, F. Shang, Y. Liu, et al., "Video super-resolution based on deep learning: a comprehensive survey," *Artificial Intelligence Review*, vol. 55, pp. 5981-6035, 2022.
- [8] S. Anwar, S. Khan, and N. Barnes, "A deep journey into super-resolution: A survey," *ACM Computing Surveys (CSUR)*, vol. 53, pp. 1-34, 2020.
- [9] C. Qiao, D. Li, Y. Liu, S. Zhang, K. Liu, C. Liu, et al., "Rationalized deep learning super-resolution microscopy for sustained live imaging of rapid subcellular processes," *Nature biotechnology*, vol. 41, pp. 367-377, 2023.
- [10] M. Chaika, S. Afat, D. Wessling, C. Afat, D. Nickel, S. Kannengiesser, et al., "Deep learning-based super-resolution gradient echo imaging of the pancreas: Improvement of image quality and reduction of acquisition time," *Diagnostic and Interventional Imaging*, vol. 104, pp. 53-59, 2023.
- [11] B. Niu, Q. Feng, J. Yang, B. Chen, B. Gao, J. Liu, et al., "Solid waste mapping based on very high resolution remote sensing imagery and a novel deep learning approach," *Geocarto International*, vol. 38, p. 2164361, 2023.
- [12] G. Liang, U. KinTak, H. Yin, J. Liu, and H. Luo, "Multi-scale hybrid attention graph convolution neural network for remote sensing images super-resolution," *Signal Processing*, vol. 207, p. 108954, 2023.
- [13] W. Yang, X. Zhang, Y. Tian, W. Wang, J.-H. Xue, and Q. Liao, "Deep learning for single image super-resolution: A brief review," *IEEE Transactions on Multimedia*, vol. 21, pp. 3106-3121, 2019.
- [14] Y. Li, B. Sixou, and F. Peyrin, "A review of the deep learning methods for medical images super-resolution problems," *Irbm*, vol. 42, pp. 120-133, 2021.
- [15] Y. LeCun, "LeNet-5, convolutional neural networks," URL: <http://yann.lecun.com/exdb/lenet>, vol. 20, p. 14, 2015.
- [16] Y. Luo, L. Zhou, S. Wang, and Z. Wang, "Video satellite imagery super-resolution via convolutional



- [32] Y. Liu, D. Yang, F. Zhang, Q. Xie, and C. Zhang, "Deep recurrent residual channel attention network for single image super-resolution," *The Visual Computer*, pp. 1-16, 2023.
- [33] W. Weng, Y. Zhang, and Z. Xiong, "Boosting event stream super-resolution with a recurrent neural network," in *European Conference on Computer Vision*, 2022, pp. 470-488.
- [34] W. Han, S. Chang, D. Liu, M. Yu, M. Witbrock, and T. S. Huang, "Image super-resolution via dual-state recurrent networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 1654-1663.
- [35] Y. Fu, Z. Liang, and S. You, "Bidirectional 3D quasi-recurrent neural network for hyperspectral image super-resolution," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 2674-2688, 2021.
- [36] M. R. Arefin, V. Michalski, P.-L. St-Charles, A. Kalaitzis, S. Kim, S. E. Kahou, et al., "Multi-image super-resolution for remote sensing using deep recurrent networks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2020, pp. 206-207.
- [37] J. Kim, J. K. Lee, and K. M. Lee, "Deeply-recursive convolutional network for image super-resolution," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 1637-1645.
- [38] W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang, "Fast and accurate image super-resolution with deep laplacian pyramid networks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 41, pp. 2599-2613, 2018.
- [39] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, pp. 1735-1780, 1997.
- [40] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space odyssey," *IEEE transactions on neural networks and learning systems*, vol. 28, pp. 2222-2232, 2016.
- [41] Y. Chang and B. Luo, "Bidirectional convolutional LSTM neural network for remote sensing image super-resolution," *Remote Sensing*, vol. 11, p. 2333, 2019.
- [42] H. Zhu, M. Guo, H. Li, Q. Wang, and A. Robles-Kelly, "Breaking the spatio-angular trade-off for light field super-resolution via lstm modelling on epipolar plane images," *arXiv preprint arXiv:1902.05672*, 2019.
- [43] C. Chou, J. Park, and E. Chou, "Generating high-resolution climate change projections using super-resolution convolutional LSTM neural networks," in *2021 13th International Conference on Advanced Computational Intelligence (ICACI)*, 2021, pp. 293-298.
- [44] X. Lu, X. Liu, L. Zhang, F. Jia, and Y. Yang, "Hyperspectral image super-resolution based on attention ConvBiLSTM network," *International Journal of Remote Sensing*, vol. 43, pp. 5059-5074, 2022.
- [45] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.
- [46] T. Lu, J. Wang, Y. Zhang, Z. Wang, and J. Jiang, "Satellite image super-resolution via multi-scale residual deep neural network," *Remote Sensing*, vol. 11, p. 1588, 2019.
- neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, pp. 2398-2402, 2017.
- [17] K. Umehara, J. Ota, and T. Ishida, "Application of super-resolution convolutional neural network for enhancing image resolution in chest CT," *Journal of digital imaging*, vol. 31, pp. 441-450, 2018.
- [18] M. U. Müller, N. Ekhtiari, R. M. Almeida, and C. Rieke, "Super-resolution of multispectral satellite images using convolutional neural networks," *arXiv preprint arXiv:2002.00580*, 2020.
- [19] M. Taş and B. Yılmaz, "Super-resolution convolutional neural network based pre-processing for automatic polyp detection in colonoscopy images," *Computers & Electrical Engineering*, vol. 90, p. 106959, 2021.
- [20] H. M. Keshk and X.-C. Yin, "Satellite super-resolution images depending on deep learning methods: a comparative study," in *2017 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, 2017, pp. 1-7.
- [21] C. Tian, R. Zhuge, Z. Wu, Y. Xu, W. Zuo, C. Chen, et al., "Lightweight image super-resolution with enhanced CNN," *Knowledge-Based Systems*, vol. 205, p. 106235, 2020.
- [22] J. Yamanaka, S. Kuwashima, and T. Kurita, "Fast and accurate image super-resolution by deep CNN with skip connection and network in network," in *Neural Information Processing: 24th International Conference, ICONIP 2017, Guangzhou, China, November 14-18, 2017, Proceedings, Part II 24*, 2017, pp. 217-225.
- [23] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in *European conference on computer vision*, 2014, pp. 184-199.
- [24] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, et al., "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, 2014.
- [25] J. Zhu, C. Tan, J. Yang, G. Yang, and P. Lio', "Arbitrary Scale Super-Resolution for Medical Images," *International journal of neural systems*, vol. 31, p. 2150037, 2021.
- [26] Y. Gu, Z. Zeng, H. Chen, J. Wei, Y. Zhang, B. Chen, et al., "MedSRGAN: medical images super-resolution using generative adversarial networks," *Multimedia Tools and Applications*, vol. 79, pp. 21815-21840, 2020.
- [27] H. Liu, J. Liu, S. Hou, T. Tao, and J. Han, "Perception consistency ultrasound image super-resolution via self-supervised CycleGAN," *Neural Computing and Applications*, pp. 1-11, 2021.
- [28] D. Mahapatra, B. Bozorgtabar, and R. Garnavi, "Image super-resolution using progressive generative adversarial networks for medical image analysis," *Computerized Medical Imaging and Graphics*, vol. 71, pp. 30-39, 2019.
- [29] X. Bing, W. Zhang, L. Zheng, and Y. Zhang, "Medical image super-resolution using improved generative adversarial networks," *IEEE Access*, vol. 7, pp. 145030-145038, 2019.
- [30] X. Yu and F. Porikli, "Ultra-resolving face images by discriminative generative networks," in *European conference on computer vision*, 2016, pp. 318-333.
- [31] L. R. Medsker and L. Jain, "Recurrent neural networks," *Design and Applications*, vol. 5, p. 2, 2001.



- [61] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 586-595.
- [62] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a "completely blind" image quality analyzer," *IEEE Signal processing letters*, vol. 20, pp. 209-212, 2012.
- [63] C. Ma, C.-Y. Yang, X. Yang, and M.-H. Yang, "Learning a no-reference quality metric for single-image super-resolution," *Computer Vision and Image Understanding*, vol. 158, pp. 1-16, 2017.
- [64] N. Venkatanath, D. Praneeth, M. C. Bh, S. S. Channappayya, and S. S. Medasani, "Blind image quality evaluation using perception based features," in *2015 twenty first national conference on communications (NCC)*, 2015, pp. 1-6.
- [65] L. Hitachi, "Super-resolution technology to convert video of various resolutions to high-definition," ed.
- [66] Y. Wang, R. Fevig, and R. R. Schultz, "Super-resolution mosaicking of UAV surveillance video," in *2008 15th IEEE International Conference on Image Processing*, 2008, pp. 345-348.
- [67] M. D. Robinson, S. J. Chiu, C. A. Toth, J. A. Izatt, J. Y. Lo, and S. Farsiu, "New applications of super-resolution in medical imaging," in *Super-resolution imaging*, ed: CRC Press, 2017, pp. 383-412.
- [68] R. Willett, I. Jermyn, R. Nowak, and J. Zerubia, "Wavelet-based superresolution in astronomy," ed: Astronomical Society of the Pacific, 2003.
- [69] W. Liu, D. Lin, and X. Tang, "Hallucinating faces: Tensorpatch super-resolution and coupled residue compensation," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 2005, pp. 478-484.
- [70] Z. Yuan, J. Wu, S.-i. Kamata, A. Ahrary, and P. Yan, "Fingerprint image enhancement by super-resolution with early stopping," in *2009 IEEE International Conference on Intelligent Computing and Intelligent Systems*, 2009, pp. 527-531.
- [71] G. Fahmy, "Super-resolution construction of iris images from a visual low resolution face video," in *2007 9th International Symposium on Signal Processing and Its Applications*, 2007, pp. 1-4.
- [47] W. Muhammad, Z. Bhutto, S. Masroor, M. H. Shaikh, J. Shah, and A. Hussain, "IRMIRS: Inception-ResNet-Based Network for MRI Image Super-Resolution," *CMES-Computer Modeling in Engineering & Sciences*, vol. 136, 2023.
- [48] Q. Qin, J. Dou, and Z. Tu, "Deep ResNet based remote sensing image super-resolution reconstruction in discrete wavelet domain," *Pattern Recognition and Image Analysis*, vol. 30, pp. 541-550, 2020.
- [49] D.-W. Jang and R.-H. Park, "Densenet with deep residual channel-attention blocks for single image super resolution," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2019, pp. 0-0.
- [50] S. Bell-Kligler, A. Shocher, and M. Irani, "Blind super-resolution kernel estimation using an internal-gan," *Advances in Neural Information Processing Systems*, vol. 32, 2019.
- [51] E. Agustsson and R. Timofte, "Ntire 2017 challenge on single image super-resolution: Dataset and study," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2017, pp. 126-135.
- [52] J. Cai, H. Zeng, H. Yong, Z. Cao, and L. Zhang, "Toward real-world single image super-resolution: A new benchmark and a new model," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 3086-3095.
- [53] P. Wei, Z. Xie, H. Lu, Z. Zhan, Q. Ye, W. Zuo, et al., "Component divide-and-conquer for real-world image super-resolution," in *Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VIII 16*, 2020, pp. 101-117.
- [54] C. Chen, Z. Xiong, X. Tian, Z.-J. Zha, and F. Wu, "Camera lens super-resolution," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 1652-1660.
- [55] X. Zhang, Q. Chen, R. Ng, and V. Koltun, "Zoom to learn, learn to zoom," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 3762-3770.
- [56] W. Wang, E. Xie, X. Liu, W. Wang, D. Liang, C. Shen, et al., "Scene text image super-resolution in the wild," in *Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part X 16*, 2020, pp. 650-666.
- [57] T. Köhler, M. Bätz, F. Naderi, A. Kaup, A. Maier, and C. Riess, "Toward bridging the simulated-to-real gap: Benchmarking super-resolution on real data," *IEEE transactions on pattern analysis and machine intelligence*, vol. 42, pp. 2944-2959, 2019.
- [58] H. R. V. Joze, I. Zharkov, K. Powell, C. Ringler, L. Liang, A. Roulston, et al., "Imagepairs: Realistic super-resolution dataset via beam splitter camera rig," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2020, pp. 518-519.
- [59] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, pp. 600-612, 2004.
- [60] H. R. Sheikh, A. C. Bovik, and G. De Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics," *IEEE Transactions on image processing*, vol. 14, pp. 2117-2128, 2005.

