

به نام خداوند جان و خرد



وزارت علوم، تحقیقات و فناوری

دانشگاه تهران

دانشکده مهندسی عمران

پایان نامه کارشناسی ارشد در رشته مهندسی نقشه برداری-گرایش فتوگرامتری

عنوان

## **بهبود کیفیت مکانی تصاویر ماهواره ای با وضوح مکانی متوسط از طریق روش های یادگیری عمیق**

نگارش

حمید امینی امیرکلائی

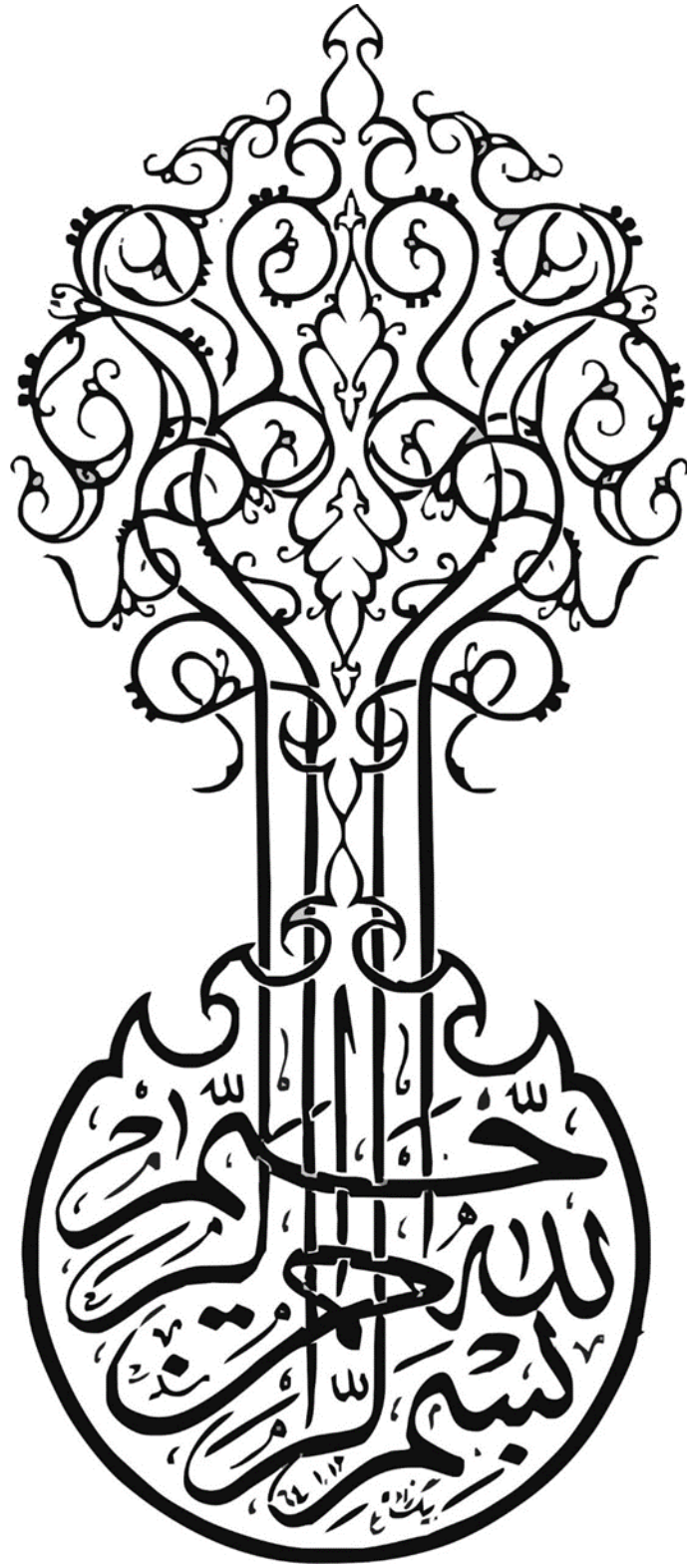
استاد راهنما

دکتر علیرضا صفدری نژاد

استاد مشاور

دکتر حامد امینی امیرکلائی

ماه و سال پایان: شهریور ۱۴۰۲



تاریخ: ۱۴۰۲/۰۸/۳۰  
شماره: ۴۹۰۲۷



دانشگاه تفرش  
اداره تحصیلات تکمیلی

### صور تجلسه دفاعیه پایان نامه کارشناسی ارشد

شماره دانشجویی: ۹۹۳۱۹۱۵۰۲

نام و نام خانوادگی: حمید امینی امیرکلاتی

رشته تحصیلی/گرایش: مهندسی نقشه برداری - فتوگرامتری

عنوان پروژه: بهبود کیفیت مکانی تصاویر ماهواره ای با وضوح مکانی متوسط از طریق روش های یادگیری عمیق

تاریخ دفاع: ۱۴۰۲/۰۶/۲۹

تاریخ تصویب: ۱۴۰۰/۱۲/۲۳

تعداد واحد: ۶

درجه:

عالی (۱۹-۲۰)  خیلی خوب (۱۸-۱۸/۹۹)  خوب (۱۶-۱۷/۹۹)  متوسط (۱۴-۱۵/۹۹)  مردود (کمتر از ۱۴)

نام و نام خانوادگی	سمت	رتبه	محل اشتغال	محل امضاء
دکتر علیرضا صفدری نژاد	استاد راهنمای اول	استادیار	دانشگاه تفرش	
-	استاد راهنمای دوم	-	-	-
دکتر حامد امینی امیرکلاتی	استاد مشاور	-	-	غایب
دکتر سیدعبدالله کیانزاد	داور داخلی	استادیار	دانشگاه تفرش	
دکتر علی محمد فتوحی	داور خارجی	استادیار	دانشگاه تفرش	
دکتر سیدعبدالله کیانزاد	نماینده تحصیلات تکمیلی	استادیار	دانشگاه تفرش	

رئیس دانشکده: دکتر بهرام بنیانیان

امضاء و مهر:

تاریخ:

مدیر تحصیلات تکمیلی دانشگاه: دکتر محمدجواد عبدالهی فرد

امضاء و مهر:

تاریخ:

به نام خدا



اینجانب حمید امینی امیرکلائی به شماره دانشجویی ۹۹۳۱۹۱۵۰۲ دانش‌آموخته رشته-گرایش مهندسی نقشه برداری-فتوگرامتری که در تاریخ ۱۴۰۲/۰۷/۲۴ از پایان نامه خود دفاع کرده‌ام، متعهد می‌شوم که مطالب مندرج در این سند علمی، برونداد کار پژوهشی اینجانب زیر نظر و راهنمایی اساتید دانشگاه تفرش بوده و به دستاوردهای دیگران که در این پژوهش از آن‌ها استفاده شده است طبق مقررات و روال متعارف ارجاع داده و مشخصات آن‌ها در فهرست منابع آورده شده است. این پایان‌نامه، پیش از این برای احراز هیچ مدرک هم‌سطح یا بالاتر ارائه نشده است. در صورت اثبات تخلف در هر زمان، مدرک تحصیلی صادر شده توسط دانشگاه از درجه اعتبار ساقط بوده و دانشگاه حق پیگیری قانونی خواهد داشت.

همه نتایج و حقوق حاصل از این پژوهش و سند متعلق به دانشگاه تفرش است و هر گونه استفاده از نتایج علمی و عملی آن، از جمله واگذاری اطلاعات به دیگران، چاپ و تکثیر، نسخه‌برداری، ترجمه و اقتباس، بدون موافقت کتبی دانشگاه تفرش ممنوع است. نقل مطالب با ارجاع درست به منبع مجاز است.

نام و نام خانوادگی دانشجو: حمید امینی امیرکلائی

تاریخ: ۱۴۰۲/۰۸/۱۱

امضاء



حمید امینی امیرکلائی

تقدیم به

مقدس‌ترین واژه‌ها در لغت نامه دلم، مادر مهربانم که زندگی‌ام را مدیون مهر و

عطوفت آن می‌دانم و پدرم

که همواره در سایه‌سار حمایتش، جرئت چشم دوختن به افق‌های دور را یافته‌ام.

## تقدیر و تشکر

اکنون که با همراهی پروردگار و پس از زحمات بسیار نگارش این تحقیق به اتمام رسیده است، بر خود وظیفه می‌دانم که از زحمات تمام افرادی که در این مسیر مرا همراهی نمودند تشکر و قدردانی نمایم.

قطعاً، اتمام این تحقیق تنها به لطف پروردگار بزرگ و همچنین راهنمایی و همراهی‌های استاد گرانقدر جناب آقای دکتر علیرضا صفدری نژاد و برادرم عزیزم دکتر حامد امینی امیرکلایی میسر گردیده است. به پاس از زحمات بی‌دریغشان که در کمال خوش‌رویی و حسن خلق بارانمایی‌ها و نظرات گران‌بهایشان در اتمام این پژوهش مرا یاری نمودند، از ایشان تشکر و قدردانی می‌نمایم.

از هیئت‌داوران جناب آقایان **دکتر کیانژاد** و **دکتر فتوحی** که با نظرات خود موجب بهبود این پژوهش گشتند، تشکر و قدردانی می‌نمایم.

در پایان از خانواده عزیزم به‌خصوص پدر زحمت‌کش و مادر مهربانم که همواره با حمایت‌ها و پشتیبانی‌های توصیف‌ناپذیرشان در این راه مرا یاری نمودند قدردانی می‌نمایم.

## چکیده

بهبود وضوح مکانی تصاویر چند طیفی یکی از پیش‌پردازش‌های رایج در استفاده حداکثری از ظرفیت این تصاویر در کاربردهای سنجش‌ازدوری است. وجود تصاویر با وضوح مکانی بهتر در کنار تصاویر چند طیفی امکان انجام این فرایندها را از طریق روش‌های تلفیق تصاویر فراهم می‌آورد. عدم اخذ هم‌زمان تصاویر با وضوح مکانی بالاتر در یک سکوی فضایی چالش‌های مربوط به هم‌رنگ‌سازی و عدم هم‌زمانی را هنگام به‌کارگیری تصاویر سایر سنجنده‌های فضایی به فرایند تلفیق تصاویر تحمیل می‌سازد. در چنین شرایطی، روش‌های تولید تصاویر به کمک سوپرزولوشن رویکردهای جایگزین برای افزایش وضوح مکانی محسوب می‌شوند. استفاده از شبکه مولد تخصصی یکی از روش‌های مؤثر در این حوزه بوده که آموزش عمیق آن‌ها نیازمند وجود داده‌های آموزشی متعدد است. عموماً دسترسی به دو نسخه تصویر فضایی با توان تفکیک طیفی مشابه و وضوح مکانی متفاوت برای آموزش چنین شبکه‌هایی میسر نیست. از این‌رو، در این تحقیق روند بهبود وضوح مکانی تصاویر چند طیفی از طریق روشی دومارحله‌ای به اجرا رسیده است. در گام نخست از یک شبکه عصبی کانولوشنی برای تولید نسخه واضح‌تر از ترکیب رنگی مرئی از تصاویر چند طیفی استفاده شده که آموزش تقویتی آن به کمک تصاویر بازآرایی شده از بستر گوگل-ارث تأمین می‌گردد. در گام دوم، تصویر با وضوح مکانی بهتر در نقش یک تصویر پان کروماتیک با تصویر چند طیفی تلفیق شده و محتوای طیفی به آن افزوده می‌شود. نتایج نشان داد که رویکرد پیشنهادی بدون نیاز به تأمین پرهزینه داده‌های آموزشی ضمن تأمین شرایط برابر در حفظ محتوای طیفی توان بهبود وضوح مکانی تصاویر چند طیفی را تا ۳۲/۸۵ درصد بهتر از بهترین روش مقایسه‌ای دارد.

## واژه‌های کلیدی

تلفیق تصاویر، سوپرزولوشن، تصاویر چند طیفی، یادگیری عمیق، آموزش تقویتی

## فهرست مطالب

پیشگفتار.....	۱۷
فصل ۱: درباره مساله.....	۱۸
۱-۱- مقدمه و تعریف مساله.....	۱۸
۱-۲-۱- سؤالات تحقیق.....	۲۱
۱-۳-۱- اهداف تحقیق.....	۲۲
فصل ۲: مروری بر پیشینه و روش های حل مساله.....	۲۴
۱-۲- مقدمه.....	۲۴
۲-۲- تحقیقات پیشین مربوط به روش های تلفیق.....	۲۶
۱-۲-۲- مروری کوتاه بر مفاهیم پایه روش های تلفیق.....	۲۶
۱-۲-۲-۱- روش های جایگزینی مؤلفه.....	۲۷
۱-۲-۲-۲- روش های تجزیه و تحلیل چند مقیاسی.....	۳۰
۱-۲-۲-۳- روش های ترکیبی.....	۳۲
۱-۲-۲-۴- روش های بیزی و فاکتورهای مبتنی بر ماتریس.....	۳۳
۱-۲-۲-۲- برخی از تحقیقات انجام شده مبتنی بر روش های تلفیق.....	۳۴
۱-۳-۲-۳- تحقیقات پیشین مربوط به روش های سوپررزولوشن.....	۴۰
۱-۳-۲-۱- روش های مبتنی بر درون یابی.....	۴۰



- ۴۲..... ۲-۳-۲- روش‌های مبتنی بر یادگیری
- ۴۲..... ۲-۳-۲-۱- مروری کوتاه بر مفاهیم پایه روش‌های سوپرزولوشن
- ۴۳..... ۲-۳-۲-۱-۱- شبکه‌های عصبی مصنوعی
- ۴۷..... ۲-۳-۲-۱-۲- لایه‌های کانولوشنی
- ۵۱..... ۲-۳-۲-۱-۳- لایه ادغام
- ۵۳..... ۲-۳-۲-۱-۴- لایه تماماًمتصل
- ۵۴..... ۲-۳-۲-۱-۵- توابع فعال‌سازی
- ۵۵..... ۲-۳-۲-۱-۶- مولفه‌های موثر در بهبود شبکه
- ۵۷..... ۲-۳-۲-۲- برخی از تحقیقات انجام‌شده مبتنی بر روش‌های سوپرزولوشن
- ۶۵..... ۲-۴- جمع بندی
- ۶۶..... فصل ۳: توسعه نظری راهکار پیشنهادی برای حل مساله
- ۶۶..... ۳-۱- مقدمه
- ۶۶..... ۳-۲- روش پیشنهادی
- ۶۹..... ۳-۳- شبکه عمیق GAN
- ۷۰..... ۳-۳-۱- معرفی شبکه عمیق GAN
- ۷۲..... ۳-۳-۲- معماری‌های استفاده شده در شبکه GAN پیشنهادی
- ۷۳..... ۳-۳-۲-۱- معماری ResNet
- ۷۵..... ۳-۳-۲-۲- معماری DensNet

- ۷۶.....۳-۳-۳ شبکه عمیق GAN پیشنهادی
- ۷۸.....۴-۳ آماده سازی نمونه های آموزشی برای یادگیری عمیق
- ۸۱.....۵-۳ تولید تصویر پانکروماتیک
- ۸۲.....۶-۳ تلفیق تصاویر
- ۸۳.....۷-۳ رویکردهای ارزیابی نتایج
- ۸۵.....۱-۷-۳ معیارهای طیفی
- ۸۵.....۱-۱-۷-۳ حداکثر نسبت سیگنال به نویز (PSNR)
- ۸۵.....۲-۱-۷-۳ شباهت ساختاری (SSIM)
- ۸۶.....۳-۱-۷-۳ ضریب همبستگی (CC)
- ۸۶.....۲-۷-۳ معیارهای هندسی
- ۸۶.....۱-۲-۷-۳ ضریب همبستگی مکانی (SCC)
- ۸۷.....۲-۲-۷-۳ کامل بودن لبه های تصویری (Cm) و صحت حفظ محتوای لبه (Cr)
- ۸۷.....۸-۳ جمع بندی
- ۸۹.....فصل ۴: پیاده سازی و ارزیابی نتایج
- ۸۹.....۱-۴ مقدمه
- ۹۰.....۲-۴ مناطق مطالعاتی و داده های مورد استفاده
- ۹۵.....۳-۴ پیاده سازی و ارزیابی نتایج
- ۱۰۷.....۴-۴ جمع بندی

فصل ۵: جمع بندی..... ۱۰۸

۱-۵- بحث و نتیجه گیری..... ۱۰۸

۲-۵- پیشنهادها..... ۱۰۹

۳-۵- قدر دانی و تشکر..... ۱۱۰

منابع..... ۱۱۱

## فهرست شکل ها

- شکل ۱-۲: دسته بندی روش های معرفی شده در زمینه افزایش وضوح تصاویر ..... ۲۵
- شکل ۲-۲: در حقیقت پنهان کردن عبارت است از ترکیب نوارهای رنگی با وضوح پایین و نوار خاکستری با وضوح بالا و به دست آوردن تصویر رنگی با وضوح بالا..... ۲۷
- شکل ۳-۲: تصاویر به ترتیب چپ به راست و بالا به پایین. تصویر چند طیفی هایپریون، تصویر پانکروماتیک کارتوست-۱، روش های تلفیق Wavelet-IHS، FFT-IHS، Wavelet-PCA و FFT-PCA و FWavelet-PCA..... ۳۶
- شکل ۴-۲: نتایج بصری روش های مختلف تلفیق ..... ۳۸
- شکل ۵-۲: نتایج تصاویر تلفیق شده روش های مختلف ..... ۴۰
- شکل ۶-۲: نمایی از نحوه عملکرد روش های درون یابی ..... ۴۲
- شکل ۷-۲: ساختار شبکه عصبی ..... ۴۵
- شکل ۸-۲: مقایسه شماتیک شبکه عصبی پیش خورنده و بلزگشت کننده..... ۴۵
- شکل ۹-۲: شبه کد یک شبکه عصبی ..... ۴۶
- شکل ۱۰-۲: نحوه اتصال در لایه کانولوشن..... ۴۸
- شکل ۱۱-۲: نحوه اثر گذاری پدینگ بر ابعاد نقشه ویژگی، (الف) بدون پدینگ، (ب) با پدینگ..... ۴۹
- شکل ۱۲-۲: نحوه اعمال کانولوشن متورم با نرخ تورم ۲ ..... ۵۰
- شکل ۱۳-۲: عملکرد کانولوشن معکوس، (الف) یک ماتریس  $2 \times 2$  با  $S=1$  و  $P=0$ ،  $F=3$ ، (ب) یک ماتریس  $3 \times 3$  با  $S=1$ ،  $P=0$ ،  $F=3$  و  $H=1$ ..... ۵۱
- شکل ۱۴-۲: عملکرد لایه ادغام ماکزیمم ..... ۵۲
- شکل ۱۵-۲: عملکرد لایه ادغام معکوس ..... ۵۳
- شکل ۱۶-۲: نحوه عملکرد لایه تمامه متصل ..... ۵۴

- شکل ۲-۱۷: نمونه‌هایی از توابع فعال‌ساز ..... ۵۵
- شکل ۲-۱۸: (الف) پروسه حذف تصادفی در شبکه‌های معمولی، (ب) پروسه حذف تصادفی در شبکه‌های کانولوشنی ..... ۵۶
- شکل ۲-۱۹: تصویر (د) تصویر اصلی، سطر (الف) زوم ۸ برابر تصویر اصلی، سطر (ب) روش درون‌یابی spline و سطر (ج) تصاویر الگوریتم سوپرزولوشن ..... ۵۸
- شکل ۲-۲۰: ساختار شبکه عمیق کانولوشنی استفاده شده در این مقاله ..... ۵۹
- شکل ۲-۲۱: تصاویر از چپ به راست به ترتیب تصویر اورجینال، نتیجه حاصل از الگوریتم bicubic، نتایج حاصل از شبکه‌های SC، KK و نتیجه شبکه معرفی شده SRCNN می‌باشند. .... ۶۰
- شکل ۲-۲۲: معماری شبکه عمیق پیشنهادی DMCN ..... ۶۱
- شکل ۲-۲۳: تصاویر از چپ به راست به ترتیب از تصویر دوم، تصویر وضوح‌بالا، نتیجه روش bicubic، نتیجه شبکه SRCNN، نتیجه شبکه vdsr، نتیجه شبکه DMCN ..... ۶۱
- شکل ۲-۲۴: نتایج مربوط به تولید تصاویر ۶۰ متری از تصاویر ۲۰ متری. تصاویر از چپ به راست به ترتیب مربوط به تصویر ۲۰ متری واقعی، ترکیب سه باند ۶۰ متری و ترکیب سه باند ۲۰ متری تولیدشده توسط شبکه. .... ۶۲
- شکل ۲-۲۵: نتایج مربوط به تولید تصاویر ۱۰ متری از تصاویر ۲۰ متری. تصاویر از چپ به راست به ترتیب مربوط به تصویر ۱۰ متری واقعی، ترکیب سه باند ۲۰ متری و ترکیب سه باند ۱۰ متری تولیدشده توسط شبکه. .... ۶۲
- شکل ۲-۲۶: تصاویر از چپ به راست به ترتیب: تصویر HR، نتایج Bicubic، نتایج CNN-1، نتایج CNN-2، نتایج GAN-1، نتایج FGAN-2 ..... ۶۳
- شکل ۲-۲۷: نتایج ارتقا ۲ برابری وضوح. به ترتیب از چپ به راست و بالا به پایین، تصویر هدف، روش SISR، روش Bicubic، روش LLE، روش ScSR، روش LsSR، روش SINE و روش پیشنهادی. .... ۶۴

- شکل ۲-۲۸: در سمت چپ بالا تصویر سنتینل-۲ را مشاهده می‌شود و تصاویر بزرگ‌تر به ترتیب از چپ به راست، تصاویر سوپررزولوشن شده به روش bicubic، شبکه EDSR8، شبکه RCAN، شبکه SRGAN، شبکه DKN-SR-GAN می‌باشند. ۶۵.....
- شکل ۳-۱: روند نمای کلی روش پیشنهادی ۶۸.....
- شکل ۳-۲: روند نمای آموزش مولد ۷۱.....
- شکل ۳-۳: روند نمای آموزش متمایزگر ۷۲.....
- شکل ۳-۴: نحوه اتصال میانبر در معماری ResNet ۷۴.....
- شکل ۳-۵: معماری ۳۴ لایه ای شبکه ResNet ۷۴.....
- شکل ۳-۶: نحوه اتصالات در بلوک های DensNet ۷۵.....
- شکل ۳-۷: آرایش قرار گیری بلاک های DensNet در یک شبکه عمیق کلاسه بندی ۷۵.....
- شکل ۳-۸: معماری شبکه عمیق Real-ESRGAN ۷۷.....
- شکل ۳-۹: شکل سمت چپ به علت ماهیت موزائیک شده تصاویر گوگ-ارث با سنجنده های مختلف دچار تنوع و ناهمگونی رنگ شده است که در شکل وسط این چالش به کمک تناظر یابی هیستوگرام با تصویر سه بانده مرئی سنتینل-۲ کاهش یافته است. ۷۹.....
- شکل ۳-۱۰: روند شماتیک برش و باز نمونه برداری کاهشی به منظور آماده سازی داده های آموزشی از تصاویر گوگل-ارث ۸۰.....
- شکل ۳-۱۱: روند شماتیک برش موزائیکی تصاویر جهت ورود به شبکه عمیق ۸۱.....
- شکل ۳-۱۲: تولید تصویر پانکروماتیک ۸۲.....
- شکل ۳-۱۳: تلفیق تصویر پانکروماتیک تولید شده و تصویر چند طیفی ۸۳.....
- شکل ۳-۱۴: روند شماتیک تولید محتوا مورد مقایسه جهت ارزیابی میزان حفظ محتوای طیفی در روش پیشنهادی ۸۴.....

- شکل ۴-۱: روند نمای پیاده سازی روش پیشنهادی ..... ۹۰
- شکل ۴-۲: تصاویر مورد استفاده از شهر یزد و گرگان ..... ۹۳
- شکل ۴-۳: شمای نزدیک تصاویر سنجنده های مختلف از شهر یزد و گرگان. .... ۹۴
- شکل ۴-۴: نمونه برداری کاهشی تصاویر ۱ متری گوگل ارث برای آموزش شبکه در راستای بهبود وضوح مکانی تصاویر سنجنده لندست-۸ و سنتینل-۲. .... ۹۶
- شکل ۴-۵: تولید تصویر پانکروماتیک از تصاویر چند طیفی تولید شده از شبکه. در جفت تصاویر، تصویر سمت چپ تصویر چند طیفی تولید شده و تصویر سمت راست تصویر پانکروماتیک مستخرج از آن است. .... ۹۷
- شکل ۴-۶: نمایی از نتایج کسب شده از روش پیشنهادی و سایر روش های مقایسه ای در مورد تصاویر سنجنده سنتینل-۲؛ شهر یزد. .... ۹۹
- شکل ۴-۷: نمایی از نتایج کسب شده از روش پیشنهادی و سایر روش های مقایسه ای در مورد تصاویر سنجنده سنتینل-۲؛ شهر گرگان. .... ۱۰۰
- شکل ۴-۸: نمایی از نتایج کسب شده از روش پیشنهادی و سایر روش های مقایسه ای در مورد تصاویر سنجنده لندست-۸؛ شهر یزد. .... ۱۰۱
- شکل ۴-۹: نمایی از نتایج کسب شده از روش پیشنهادی و سایر روش های مقایسه ای در مورد تصاویر سنجنده لندست-۸؛ شهر گرگان. .... ۱۰۲
- شکل ۴-۱۰: مقایسه بصری تصاویر گرادیان بمنظور ارزیابی بصری میزان حفظ محتوای مکانی در روش های پیشنهادی و مقایسه ای مربوط به سنجنده سنتینل-۲؛ دو سطر بالا مجموعه داده شهر گرگان؛ دو سطر پایین مجموعه داده شهر یزد ..... ۱۰۳
- شکل ۴-۱۱: مقایسه بصری تصاویر گرادیان بمنظور ارزیابی بصری میزان حفظ محتوای مکانی در روش های پیشنهادی و مقایسه ای مربوط به سنجنده لندست-۸؛ دو سطر بالا مجموعه داده شهر گرگان؛ دو سطر پایین مجموعه داده شهر یزد ..... ۱۰۴

## فهرست جداول

- جدول ۴-۱: معرفی باندهای طیفی سنجنده لندست-۸..... ۹۱
- جدول ۴-۲: معرفی باندهای طیفی سنجنده سنتینل-۲..... ۹۲
- جدول ۴-۳: جزئیات مربوط به منطقه مورد مطالعه در ارزیابی نتایج روش پیشنهادی ..... ۹۴
- جدول ۴-۴: روش و تنظیمات مرتبط با روند آموزش تقویتی شبکه عصبی Real-ESRGAN..... ۹۷
- جدول ۴-۵: تعداد داده های آموزشی استفاده شده در این تحقیق..... ۹۸
- جدول ۴-۶: مقایسه کمی نتایج روش پیشنهادی و روش های مقایسه ای از منظر میزان حفظ محتوای طیفی و مکانی..... ۱۰۵



پایان نامه ارائه شده در ۶ فصل تهیه شده است. فصل اول نامه کلیاتی از موضوع مورد نظر که بهبود کیفیت مکانی تصاویر ماهواره ای با وضوح مکانی متوسط از طریق روش‌های یادگیری عمیق می‌باشد، ارائه شده است. در فصل دوم روش‌های پیشین سنتی و پیشرفته در زمینه بهبود وضوح تصاویر سنجش از دوری به صورت جداگانه و مفصل تشریح گشته است و برای درک بهتر روش‌های معرفی شده در هر بخش مروری کوتاه بر مفاهیم پایه ارائه شده است. در فصل سوم به طور مفصل روش پیشنهادی تشریح شده و الگوریتم‌های ارزیابی نتایج معرفی شده اند، به طوری که در معرفی روش پیشنهادی جزئیات شبکه عمیق از جمله لایه‌ها، نحوه آموزش، روند استفاده از شبکه آموزش دیده و در آخر نحوه تلفیق تصاویر به طور کامل تشریح شده است. فصل چهارم به معرفی منطقه مورد مطالعه، داده‌های مورد استفاده، پیاده سازی روش پیشنهادی به کمک داده‌ها و ارزیابی نتایج به صورت بصری و عددی در قالب جداول و شکل‌ها پرداخته شد. فصل پنجم حاوی نتایج بدست آمده و پیشنهادات آتی مطرح شده، می‌باشد و در نهایت فصل ششم منابع مورد استفاده در این تحقیق ارائه شده اند.

## فصل ۱: درباره مساله

### ۱-۱- مقدمه و تعریف مساله

امروزه برای بسیاری از کاربردها در حوزه مدیریت زمین مانند پایش تغییرات، طبقه‌بندی پوشش‌های زمینی، کاربری اراضی، برنامه‌ریزی شهری و غیره اغلب استفاده از تصاویر سنجش از دوری با قدرت تفکیک طیفی و مکانی بالا را می‌طلبند [۱]. اما همواره این‌گونه تصاویر در دسترس نیستند. تصاویر هوایی، دارای قدرت تفکیک مکانی بالایی هستند ولی اخذ این تصاویر چالش‌هایی همچون هزینه بسیار بالای پرواز و شرایط آب و هوایی خاص را در خود دارد. با پیشرفت علم و فناوری و ظهور ماهواره‌ها، این امکان به وجود آمد که از روی سکوه‌های ماهواره‌ای نیز تصویربرداری از زمین صورت گیرد. با گسترش تعداد ماهواره‌های متنوع و پیشرفته شدن عدسی‌های نصب‌شده بر روی سکوه‌های فضایی این امکان به وجود آمد که تصاویر با توان تفکیک مکانی و طیفی متفاوت اخذ گردد. به‌عنوان مثال سنجنده وردیو-۲<sup>۱</sup> توانایی ثبت تصاویر پانکروماتیک با وضوح ۰/۴۶ متری و ثبت تصاویر چند طیفی دارای ۸ باند با وضوح مکانی ۱/۲ متری را دارند [۲]. با این حال برای استفاده از تصاویر ماهواره‌ای با قدرت تفکیک مکانی بالا لازم است این تصاویر خریداری شوند و به‌صورت رایگان در دسترس عموم قرار ندارند. بنابراین برای پوشش یک منطقه بزرگ و تجزیه و تحلیل در راستای اهداف مختلف بر روی آن نیازمند صرف هزینه بسیار زیادی جهت اخذ تصویر است. داده‌های گوگل-ارث<sup>۲</sup> متشکل از تصاویر سنجنده‌های با وضوح مکانی بالا هستند، اما این تصاویر به‌روز و قابل دسترس در هر هفته نیستند و به علت داشتن تنها سه باند رنگی که از نظر طیفی دقیق نیستند،

---

<sup>۱</sup> Worldview-2

<sup>۲</sup> GoogleEarth

کارایی آن‌ها محدود می‌باشد. در این میان سنجنده هایی با وضوح مکانی متوسط مانند سنجنده لندست<sup>۱</sup>-۸ و سنتینل<sup>۲</sup>-۲ وجود دارند که داده‌های آن‌ها به صورت رایگان در اختیار عموم قرار دارند و همچنین، دوره زمانی اخذ تصویر بسیار پایینی دارند. اما این سنجنده‌ها دارای توان تفکیک مکانی مطلوبی در زمینه های شناسایی ساختمان، شناسایی تغییرات، ارزیابی تخریب و ... نمی‌باشد [۳]. با این حال این سنجنده ها به علت اطلاعات طیفی مطلوبی که دارند در حوزه‌هایی همچون پوشش زمین، کشاورزی، جنگل‌داری و نظارت بر بخار آب و ابرها کاربردی هستند و همچنین از مزایای در دسترس بودن، دوره زمانی کوتاه اخذ تصویر و رایگان بودن برخوردارند که بنابراین موضوع، لازم است توجه بیشتری به این داده‌ها صورت گیرد [۴]. با توجه به محدودیت‌ها و چالش‌های مطرح شده در تهیه تصاویر ماهواره‌ای، در این پژوهش تلاش شده تاراهکاری در راستای بهبود وضوح مکانی تصاویر سنجش از دوری در عین حفظ اطلاعات طیفی ارائه شود.

امروزه سامانه‌های سنجش از دوری متعددی به منظور پایش زمین با توان تفکیک مکانی، طیفی و زمانی مختلف توسعه یافته‌اند. از نظرگاه فنی، توان تفکیک مکانی و طیفی سنجنده‌های فضایی ارتباط معکوسی را برخوردارند. به عبارت بهتر هرچه توان تفکیک طیفی سنجنده ای بیشتر باشد توان تفکیک مکانی آن کمتر خواهد بود [۵]. علیرغم گستردگی کاربردهای تصاویر سنجش از دوری با وضوح طیفی بالا، وضوح مکانی متوسط و یا کم آن‌ها همواره به عنوان یک عامل محدودکننده در برخی از کاربردها مانند شناسایی ساختمانی، شناسایی تخریب، گسترش شهری و غیره به شمار می‌رود. روش‌های محاسباتی مختلفی به منظور بهبود توان تفکیک مکانی تصاویر چند طیفی توسعه یافته‌اند که عموماً به یک تصویر با توان تفکیک مکانی بالاتر (تصویر پانکروماتیک<sup>۳</sup>) نیازمندند. در برخی از سکوه‌های فضایی امکان اخذ هم‌زمان تصاویر پانکروماتیک با توان تفکیک مکانی بالاتر در کنار تصاویر چند طیفی وجود داشته و این موضوع امکان تلفیق آن‌ها را به سهولت فراهم می‌آورد. با این حال، برخی از سنجنده‌های

---

Landsat<sup>۱</sup>

Sentinel<sup>۲</sup>

Panchromatic<sup>۳</sup>

فضایی فاقد چنین ترکیبی از تصاویر چندطیفی و پانکروماتیک بوده و بهبود توان تفکیک مکانی آن‌ها نیازمند دسترسی به تصاویر با توان تفکیک مکانی بالاتر از سایر منابع است که این امر موجب مشکلات و محدودیت‌هایی خواهد شد. محدودیت‌هایی از قبیل عدم هم‌زمانی تصاویر، هم مرجع نبودن تصاویر، تمایز زاویه تصویربرداری و تمایز در چگونگی رخداد اعوجاجات هندسی چالش‌های اجرای روند تلفیق تصاویر در هنگام به‌کارگیری تصاویر اخذشده از منابع متفاوت است.

سازوکار روش‌های تلفیق که در اصطلاح رایج‌تر، روش‌های پن‌شارپنینگ<sup>۱</sup> نام‌گذاری می‌شوند، بر پایه بیشینه‌سازی هم‌زمان مزایای طیفی تصاویر چند طیفی و مکانی تصاویر پانکروماتیک در تولید محصول نهایی نباشده است. به عبارت بهتر استفاده هم‌زمان از مزیت وضوح مکانی تصویر پانکروماتیک و وضوح طیفی تصویر چند طیفی اساس روش‌های تلفیق می‌باشد. محصول این روش‌ها، تصاویر با توان تفکیک مکانی مشابه با تصویر پانکروماتیک و محتوای طیفی شبیه به تصاویر چند طیفی است. در روش‌های تلفیق تصویر، بهبود توان تفکیک مکانی نیازمند دسترسی هم‌زمان به تصویر با توان تفکیک مکانی بالا (تصویر پانکروماتیک) و تصویر چند طیفی متناظر بوده که تأمین آن همیشه ممکن نبوده و این موضوع همواره ضعف این روش‌ها به شمار می‌آید. روش‌هایی بانام سوپرزولوشن<sup>۲</sup> یا تقویت وضوح برای چنین شرایطی توسعه‌یافته‌اند. ایده اصلی این روش‌ها تولید تصاویر با توان تفکیک مکانی بهتر در زمان فقدان محتوای اطلاعاتی با وضوح بالاتر است. موضوع توسعه روش‌های سوپر رزولوشن اولین بار در سال ۱۹۸۰ مورد مطالعه قرار گرفت و پس‌از آن روش‌های گوناگونی در چهار دهه گذشته توسعه‌یافته است. در این روش‌ها با در اختیار داشتن تعداد بسیار زیادی از تصاویر آموزشی متناظر در دو نسخه HR<sup>۳</sup> و LR<sup>۴</sup>، تلاش می‌شود رابطه‌ای بین این تصاویر از طریق روش‌های یادگیری برقرار گردد. به عبارت بهتر با داشتن تصویر LR و تصویر مورد انتظار HR مدلی به دست می‌آید که توانایی تبدیل تصویر LR به تصویر HR را دارد و در نهایت

---

<sup>۱</sup> Pan-Sharpning

<sup>۲</sup> Super Resolution

<sup>۳</sup> High Resolution

<sup>۴</sup> Low Resolution

مدل آموزش یافته برای افزایش وضوح تصویر LR استفاده می‌شوند [۶]. برخی از روش‌های پیشرفته به کاررفته در این حوزه عبارت‌اند از: [7](GAN<sup>1</sup>), [8](SRCNN<sup>2</sup>), [9](BSRGAN<sup>3</sup>), [4](CDC), [10](DRCN<sup>5</sup>), [11](ESPCN) و [11](SESR<sup>6</sup>). روش‌های مبتنی بر یادگیری به علت پردازش‌های تودرتو و نرمال‌سازی‌های متعدد و حتی در برخی موارد تغییر فرمت تصاویر ورودی و خروجی، باعث از بین رفتن محتوای طیفی را می‌شوند.

عدم دسترسی به تصاویر HR در سنجنده‌های فاقد باند پانکروماتیک مشوقی برای به‌کارگیری روش‌های سوپررزولوشن در تولید تصاویر چند طیفی با وضوح بالاتر است. استفاده از روش‌های یادگیری عمیق در این حوزه ایده اصلی این تحقیق تلقی می‌شوند. با این حال دسترسی به تصاویر مناسب برای آموزش چنین شبکه‌هایی یک چالش اجرایی محسوب می‌شود. به عبارت بهتر، تأمین هم‌زمان تصاویر چند طیفی مناسب در نسخه HR برای به‌کارگیری در روند آموزش یک چالش اصلی در اجرای روش‌های یادگیری عمیق برای سنجنده‌های فاقد باند پانکروماتیک است. در این تحقیق راهکاری دومی ارائه شده است. این روند ضمن افزایش وضوح مکانی تصاویر چند طیفی در زمان فقدان باند طیفی پانکروماتیک ارائه شده است. این روند می‌تواند تفکیک مکانی تصاویر اولیه، محتوای طیفی را مشابه با روش‌های سنتی تلفیق تصاویر بازیابی می‌نماید. عدم نیاز به تصاویر چند طیفی با توان تفکیک مکانی بالاتر، استفاده از ظرفیت تصاویر ماهواره‌ای منتشر شده در بستر گوگل-ارث در روند یادگیری عمیق و بازیابی محتوای طیفی بر پایه روش‌های سنتی تلفیق از ویژگی‌های این روش به شمار می‌رود.

## ۱-۲- سوالات تحقیق

در روند اجرای این تحقیق سوالات زیر مطرح است:

---

<sup>1</sup> Generative Adversarial Network  
<sup>2</sup> Super-Resolution Convolutional Neural Network  
<sup>3</sup> Blind image Super-Resolution  
<sup>4</sup> Component Divide-and-Conquer  
<sup>5</sup> Deeply-Recursive Convolutional Network  
<sup>6</sup> Efficient Sub-Pixel Convolutional Neural network

- ۱- جهت ارتقا مکانی تصاویر از چه روشی استفاده شود؟
- ۲- تصاویر ارتقا یافته در مقایسه با تصاویر سنجنده های دیگر با وضوح بهتر چه مزیتی دارند؟
- ۳- جهت آموزش شبکه معرفی شده از چه تصاویری استفاده شود و آیا مقرون به صرفه است؟
- ۴- مشکل وجود تفاوت رادیومتریکی بسیار زیاد تصاویر اخذ شده از سنجنده مورد نظر و تصویر گوگل-ارث چگونه باید برطرف شود؟
- ۵- آیا نتایج تصاویر بهبود مکانی یافته اطلاعات طیفی مناسبی دارند؟ در غیر این صورت راهکار چیست؟

### ۱-۳- اهداف تحقیق

- با توجه به بخش ۱-۱ هدف نهایی در این تحقیق بهبود وضوح مکانی تصاویر چند طیفی فضایی در زمان فقدان تصاویر با توان تفکیک مکانی بالاتر به کمک روش های تلفیق تصویر و یادگیری عمیق می باشد. در این راستا اهداف زیر در روند پیشنهادی مدنظر قرار گرفته اند:
- ۱- بررسی شبکه ای عمیق با در نظر گرفتن ویژگی های متنوع در مقیاس های متفاوت برای بازیابی اطلاعات هندسی و تولید تصویر با وضوح بالاتر.
  - ۲- استفاده از تصاویر رایگان گوگل-ارث جهت آموزش شبکه عمیق
  - ۳- ارائه راهکاری جهت مشابه سازی طیفی تصاویر آموزشی و تصاویر سنجش از دوری با وضوح مکانی پایین
  - ۴- ارائه راهکاری جهت بهبود اطلاعات طیفی تصاویر ارتقا یافته
  - ۵- کاهش هزینه تولید تصاویر در کاربردهای شهری
  - ۶- تولید تصاویر با کاربردهای شهری در زمان دلخواه

## ۱-۴- ساختار پایان نامه

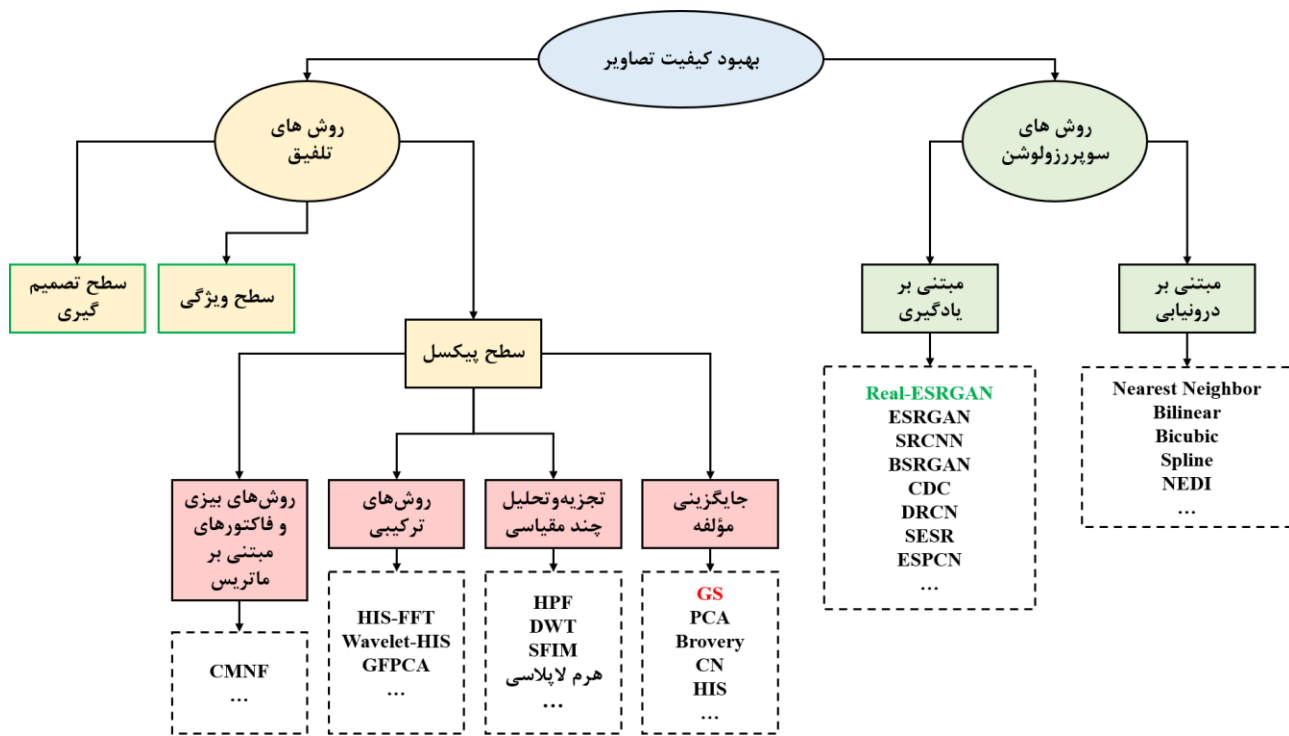
ساختار این رساله از پنج فصل تشکیل شده است. فصل اول به بررسی کلیات مطرح شده در این تحقیق شامل اهداف، ضرورتها و سوالات پرداخته است. فصل دوم از مروری جامع بر تحقیقات انجام شده و دسته بندی روش های تلفیق تصاویر تشکیل شده و در فصل سوم به معرفی روش پیشنهادی با جزئیات کامل شبکه معرفی شده، پرداخته شده است. در فصل چهارم نتایج حاصل از پیاده سازی شبکه پیشنهادی به همراه مقادیر بدست آمده از ارزیابی ها آورده شده است. در نهایت در فصل پنجم، نتایج و دستاوردهای حاصل از این تحقیق به همراه پیشنهادات آتی ارائه شده است.

## فصل ۲: مروری بر پیشینه و روش های حل مساله

### ۲-۱- مقدمه

سنجنده های متفاوت، قدرت تفکیک متفاوتی باهم دارند (قدرت تفکیک مکانی، زمانی، طیفی و رادیو متریکی). از سویی توان تفکیک مکانی و طیفی باهم رابطه عکس دارند به این منظور که هرچه توان تفکیک طیفی تصویری بیشتر باشد توان تفکیک مکانی آن کمتر می شود. تصاویر چند طیفی فضایی دارای مزایا و کاربردهای بسیاری هستند ولی وضوح مکانی ضعیف آن ها همواره یک عامل محدودکننده در این تصاویر به شمار می رود. استفاده توأم از پتانسیل اطلاعات طیفی و مکانی می تواند کیفیت نتایج قابل استخراج از داده های سنجش از دوری را ارتقا بخشد. رویکردهای بسیاری در این زمینه توسعه یافته اند که با توجه به شکل (۲-۱) در دو دسته کلی می توان آن ها را تقسیم بندی نمود: ۱- روش های تلفیق ۲- روش های سوپررزولوشن





شکل ۱-۲: دسته بندی روش های معرفی شده در زمینه افزایش وضوح تصاویر

در روش های تلفیق شرط لازم برای ارتقا وضوح مکانی، وجود تصویر پانکروماتیک است که یا از طریق همان سنجنده تولید می شود و یا از تصاویر پانکروماتیک سنجنده های موجود دیگر مورد استفاده قرار می گیرد که محدودیت های بسیاری از قبیل عدم هم زمانی تصاویر و تمایز در چگونگی رخداد اعوجاجات هندسی چالش های اجرای روند تلفیق تصاویر در هنگام به کارگیری تصاویر اخذ شده از منابع متفاوت می باشد. همچنین برای روش های تلفیق لازم است تصویر با وضوح طیفی بالا دارای هندسه کاملاً یکسانی با تصویر پانکروماتیک باشد.

از طرفی در روش های سوپررزولوشن لزومی به وجود تصویر با وضوح بالا نمی باشد و به همین علت بحث هم مرجع بودن تصویر پانکروماتیک و چند طیفی نیز مطرح نمی باشد. اما لازمه نتایج مطلوب این روش ها در وجود تعداد بسیار زیادی داده آموزشی جهت آموزش شبکه مورد نظر است. از طرفی در نتایج این روش ها به علت تغییر فرمت تصویر و ... اطلاعات طیفی تصاویر اولیه عمدتاً تخریب شده و نتایج تصاویر به لحاظ طیفی از اطلاعات واقعی تصاویر اولیه برخوردار نیستند. همچنین پردازش سنگین و زمان بر از معایب این روش ها به شمار می آید [۶].

در ادامه به بررسی پژوهش‌های پیشین در زمینه افزایش وضوح تصاویر به روش‌های تلفیق و سوپررزولوشن پرداخته شده است.

## ۲-۲- تحقیقات پیشین مربوط به روش‌های تلفیق

در ابتدای این بخش به معرفی مختصری از روش‌های تلفیق معروف پرداخته شده و سپس برخی از تحقیقات پیشین از کاربردها و روش‌های نوین تلفیق تصاویر مرور شده است.

## ۲-۲-۱- مروری کوتاه بر مفاهیم پایه روش‌های تلفیق

یکی از مهم‌ترین پردازش‌های سنجش‌ازدوری تلفیق داده می‌باشد که به‌طور کلی به‌منظور بهبود کیفیت داده و افزایش عملکرد آنالیزها انجام می‌گیرد. تلفیق تصاویر سنجش‌ازدوری در سه سطح صورت می‌پذیرد: (۱) در سطح پیکسل [۱۲]، (۲) در سطح ویژگی [۱۳]، (۳) در سطح تصمیم‌گیری [۱۴] که در این پژوهش تلفیق در سطح پیکسل انجام شده است.

یکی از مهم‌ترین و کارآمدترین تکنیک‌هایی که جهت تلفیق در سطح پیکسل ارائه شده است تکنیک پنشارپنینگ می‌باشد (شکل (۲-۲)) [۱۵]. در این تکنیک اطلاعات مکانی تصویر پانکروماتیک (که در تصویر چند طیفی وجود ندارد) و اطلاعات طیفی تصویر چند طیفی (که در مقابل تصویر تک باند پانکروماتیک است) باهم تلفیق می‌شوند و محصول نهایی، یک تصویر منحصربه‌فرد بوده که هم‌زمان اطلاعات طیفی تصویر چند طیفی و اطلاعات مکانی تصویر پانکروماتیک را در خود دارد.



شکل ۲-۲: در حقیقت پنشارپن کردن عبارت است از ترکیب نوارهای رنگی با وضوح پایین و نوار خاکستری با وضوح بالا و به دست آوردن تصویر رنگی با وضوح بالا.

به طور کلی روش های پنشارپنینگ به چهار گروه اصلی (۱) جایگزینی مؤلفه (CS<sup>۱</sup>) [16]، (۲) تجزیه و تحلیل چند مقیاسی (MRA<sup>۲</sup>) [۱۷]، (۳) روش های ترکیبی [۱۸] و (۴) روش های بیزی<sup>۳</sup> و فاکتورهای مبتنی بر ماتریس<sup>۴</sup> [۱۹]، تقسیم بندی می شوند [۲۰].

## ۲-۲-۱-۱-۲-۲ روش های جایگزینی مؤلفه

دسته اول روش های مبتنی بر جایگزینی مؤلفه هستند. در این روش ها تصویر چند طیفی که وضوح طیفی بالایی دارد، به فضای دیگری برده می شود تا اطلاعات مکانی و طیفی آن از هم جدا شود [۱۵]. سپس تصویر تبدیل شده با جایگزینی تصویر پانکروماتیک به جای اطلاعات مکانی اش، تلفیق می شود و دوباره به فضای اولیه بازمی گردد [۲۱]. در این روش ها هرچه همبستگی بین دو تصویر پانکروماتیک و تصویر جایگزین بیشتر باشد و دو تصویر به هم شبیه تر باشند، نتایج تلفیق بهتر خواهد بود. به همین منظور معمولاً پیش از فرآیند جایگزینی تصویر

<sup>۱</sup> Component Substitution

<sup>۲</sup> Multi Resolution Analysis

<sup>۳</sup> Bayesian Methods

<sup>۴</sup> Matrix Factorization

پانکروماتیک، یک انطباق هیستوگرام انجام می‌شود. در نهایت این فرآیند به کمک استفاده از یک تبدیل طیفی معکوس نتیجه نهایی حاصل می‌شود [۲۱]. این پروسه به صورت کلی بر تمام تصویر انجام می‌شود و معمولاً این دسته روش‌ها سریع و آسان هستند. از مزیت این روش‌ها می‌توان اشاره کرد که معمولاً جزئیات مکانی تصویر نهایی با کیفیت عالی مشخص می‌باشد. از طرفی این الگوریتم‌ها قادر به تشخیص اختلافات محلی بین تصویر پانکروماتیک و تصویر چند طیفی نیستند، به این علت که میان تصویر پانکروماتیک و تصویر چند طیفی اولیه عدم تطابق طیفی وجود دارد و این امر باعث به وجود آمدن اعوجاجات طیفی قابل توجهی در تصویر نهایی می‌شود. روش‌های معروف معرفی شده در این دسته شامل الگوریتم‌هایی مانند تجزیه و تحلیل مؤلفه‌های اصلی (PCA<sup>۱</sup>)، گرام اشمیت (GS<sup>۲</sup>)، الگوریتم Brovey، الگوریتم CN و اشباع شدت رنگ (IHS<sup>۳</sup>) هستند.

الگوریتم PCA: در این روش میزان اطلاعات هر باند مشخص شده و از بین آن‌ها باندی که اطلاعات بیشتری دارد مشخص می‌شود. از طرفی با کاهش تعداد باندهای تصویر چند طیفی و جایگزین شدن آن با باندهایی که اطلاعات بیشتری داشتند باعث استخراج بیشتر اطلاعات از تصاویر می‌شود. سرانجام با تبدیل اطلاعات به محورهای متعامد، تفکیک پذیری طیف‌هایی که با یکدیگر همپوشانی اطلاعات دارند افزایش می‌یابد. فرضیه اصلی این روش در کاربردهای تلفیق به این صورت است که اطلاعات مکانی در مؤلفه اصلی اول متمرکز شده است و اطلاعات طیفی در دیگر مؤلفه‌های اصلی هستند [۲۲].

الگوریتم GS: تبدیل گرام اشمیت روشی است که در جبر خطی و آمار چند متغییره استفاده می‌شود تا مجموعه‌ای از بردارها را متعامد کند. تبدیل گرام اشمیت که اغلب در پهن‌بند پهن‌بندینگ مورد استفاده قرار می‌گیرد، ابتدا توسط کداک<sup>۴</sup> پیشنهاد شد. در این روش، یک تصویر پانکروماتیک هم‌اندازه با تصویر چند طیفی از طریق میانگین‌گیری

---

Principal Component Analysis <sup>۱</sup>

Gram Schmidt <sup>۲</sup>

Intensity Hue-Saturation <sup>۳</sup>

Kodak <sup>۴</sup>

وزن دار از باندهای طیفی شبیه‌سازی شده و سایر باندهای طیفی نسبت به آن و نسبت به یکدیگر از طریق تبدیل گرام-اشمیت متعامد سازی می‌شوند. سپس باند پانکروماتیک اصلی جایگزین باند شبیه‌سازی شده و به تبع آن سایر باندهای طیفی به صورت متعامد بر آن به محاسبه می‌شوند. در انتها با اعمال تبدیل معکوس گرام-اشمیت، باندهای طیفی در ابعادی هم‌اندازه با تصویر پانکروماتیک بازسازی می‌شوند. ضرایب مربوط به بازسازی باندهای طیفی متعامد در یک‌روند تکراری و به ازای هر تصویر برآورد می‌شود [۲۳].

الگوریتم Brovey: در این روش با تبدیل Brovey باندهای تصویر چند طیفی نرمال می‌شوند و سپس با شدت و روشنایی تصویر پانکروماتیک تطبیق داده می‌شوند. ورودی و خروجی این روش تصاویری سه بانده خواهند بود (رابطه (۱)) [۲۴]

$$B = \frac{I_{MSi}}{\sum_i^n I_{MSi}} \times I_P \quad \text{رابطه (۱-۲)}$$

در رابطه (۱-۲)  $I_{MSi}$  باند  $i$  ام تصویر چند طیفی و  $I_P$  تصویر پانکروماتیک است.

الگوریتم CN: این روش بهبود داده شده روش Brovey می‌باشد که توانسته است محدودیت‌های باندی را رفع کند و این امکان را ایجاد کند که تلفیق تصویر با بیش از سه باند صورت گیرد که رابطه (۲) نحوه اجرای این روش را نشان می‌دهد. [۲۴]

$$CN = \frac{3 \times (1 + I_{MSi})(1 + I_P)}{\sum_i^n I_{MSi} + 3} \times I_P \quad \text{رابطه (۲-۲)}$$

الگوریتم IHS: تبدیل رنگی IHS در مقایسه با فضای RGB، اطلاعات مکانی و طیفی را از یکدیگر جدا می‌کند. در این فضا اطلاعات مکانی در مؤلفه شدت و اطلاعات طیفی در مؤلفه‌های اشباع و چرذگی (Hue) قرار می‌گیرند. این الگوریتم ابتدا باندهای تصویر RGB را به فضای IHS برده و سپس تصویر پانکروماتیک، با مؤلفه باند  $I$  در فضای IHS جایگزین می‌شود. در آخر تصویر تولید شده به فضای اولیه RGB بازمی‌گردد و نتیجه این فرآیند تصویری خواهد بود با اطلاعات طیفی و مکانی بالاتر [۲۵].

## ۲-۱-۲-۲- روش‌های تجزیه و تحلیل چند مقیاسی

دسته دوم از روش‌های تلفیق تصاویر، روش‌های مبتنی بر تجزیه و تحلیل چند مقیاسی است که هدف آن بکارگیری یک فیلتر مکانی بر روی تصویر پانکروماتیک و تولید جزئیات مکانی به منظور ترکیب آن با تصویر چند طیفی است. روش‌های معروف در این دسته عبارت‌اند از: الگوریتم<sup>۱</sup> HPF، الگوریتم تبدیل موجک (DWT<sup>۲</sup>)، الگوریتم روش مدولاسیون شدت مبتنی بر فیلتر هموار کننده (SFIM<sup>۳</sup>) و روش هرم لاپلاسی<sup>۴</sup>.

روش HPF: در این روش ابتدا بر روی تصویر پانکروماتیک یک فیلتر بالاگذر (۳×۳ یا ۵×۵) اعمال می‌شود که در این راستا ویژگی‌های خطی و لبه‌ها در آن بارزتر می‌شوند. سپس بر روی تصویر چند طیفی یک فیلتر پایین گذر (با ابعاد ۳×۳) اعمال می‌شود که ویژگی‌های طیفی را برجسته‌تر خواهد کرد. در پایان با توجه به رابطه (۳) تصاویر فیلتر شده باهم ترکیب می‌گردند.

$$I_{HPF} = \frac{I_{HP} + I_{HMi}}{2} \quad \text{رابطه (۳-۲)}$$

در این رابطه  $I_{HP}$  تصویر پانکروماتیک است که بر روی آن فیلتر بالاگذر اعمال شده و  $I_{HMi}$  باند  $i$  ام تصویر چند طیفی است که بر روی آن فیلتر پایین گذر اعمال شده و  $L_{HPF}$  نتیجه تلفیق به روش HPF می‌باشد [۲۶].

روش DWT: در این روش اساس بر پایه جایگذاری جزئی است. تصویر در این روش به چهار زیر تصویر یا مؤلفه‌های (LL، HL، LH و HH) به نام ضرایب موجک تبدیل می‌شود که ضریب LL زیر تصویر تقریب و ضرایب HL ضریب جزئیات

---

<sup>۱</sup> High Pass Filter

<sup>۲</sup> Discrete Wavelet Transform

<sup>۳</sup> Smoothing Filter-based Intensity Modulation

<sup>۴</sup> Laplacian Pyramid

در راستای عمودی، LH ضریب جزئیات در راستای افقی و HH ضریب جزئیات در راستای قطری می باشد که قسمت فرکانس پایین می تواند تا رسیدن به نتیجه دلخواه تجزیه شود [۲۷].

روش SFIM: در این روش به کمک نسبت بین تصویر پانکروماتیک و یک تصویر مات شده با اعمال یک فیلتر پایین گذر، تصویری با وضوح بالا و بدون از دست دادن جزئیات و کنتراست<sup>۱</sup> تولید می شود. این فرآیند به کمک اعمال یک فیلتر در حیطه مکان بر روی تصویر پانکروماتیک، جزئیات مکانی را استخراج می کند و سپس این جزئیات استخراج شده به تصویر چند طیفی اضافه می شود. بنا بر سخت بودن طراحی فیلتر مکانی، پیاده سازی این روش مشکل بوده و در مقایسه با روش های دیگر زمان بر است [۱۷].

روش هرم لاپلاسی: در این روش ابتدا به کمک اعمال فیلتر پایین گذر نسخه هایی از تصویر اولیه تولید می شود و سپس از اختلاف تصویر اصلی با نسخه های دیگرش را محاسبه می کنند که این فرآیند تجزیه هرمی نامیده می شود و اولین بار توسط آدلسون<sup>۲</sup> معرفی شد [۲۸]. اگر فیلتر استفاده شده گاوسی باشد، هرم گاوسی ایجاد خواهد شد. تفاوت بین سطوح متوالی هرم گاوسی، معرف هرم لاپلاس می باشد. در حقیقت می توان فیلترهای گاوسی را طوری تنظیم کرد تا تابع انتقال مدولاسیون (MTF<sup>۳</sup>) سنجنده را باهم مقایسه کند. در این راستا، پارامتری که باعث تشخیص کل توزیع می شود، انحراف استاندارد گاوسی است که از اطلاعات مبتنی بر سنسور تعیین می شود. بنابراین، دو سطح تزریق مورد استفاده قرار گرفته است که به ترتیب MTF-GLP<sup>۴</sup> و MTF-GLP-HPM<sup>۵</sup> می باشند [۱۲].

---

Contrast<sup>۱</sup>

Burt and Adelson<sup>۲</sup>

Modulation Transfer Function<sup>۳</sup>

MTF-Generalized Laplacian Pyramid<sup>۴</sup>

MTF-GLP with High Pass Modulation<sup>۵</sup>

## ۲-۱-۳- روش‌های ترکیبی

دسته سوم از روش‌های تلفیق روش‌های ترکیبی می‌باشد که در این روش‌ها ترکیبی از روش‌های دسته اول، روش‌های دسته دوم و یا ترکیبی از هر دو مورد استفاده قرار می‌گیرد. برخی از روش‌های معرفی شده در این دسته عبارت‌اند از: روش<sup>۱</sup> IHS-FFT، روش Wavelet-IHS و روش فیلتر گریدی تحلیل مؤلفه اصلی (GFPCA<sup>۲</sup>).

روش IHS-FFT: در سال ۲۰۰۷ این روش توسط لینگ<sup>۳</sup> معرفی شد که روش پیشرفته‌تری نسبت به روش قدیمی IHS بود و باعث افزایش دقت تلفیق و کاهش اعوجاجات طیفی شده است. مهم‌ترین ایده این روش جایگزینی جزئی تصویر پانکروماتیک در جای مؤلفه شدت است که در این راستا از روش فیلترینگ فوریه سریع (FFT) استفاده می‌شود. باقی مراحل مانند روش IHS انجام می‌شود [۲۹].

روش Wavelet-IHS: این روش در ابتدا در سال ۲۰۰۵ توسط ژانگ<sup>۴</sup> معرفی شد که هدف آن کاهش اعوجاجات طیفی ناشی از تلفیق تصاویر بود. همانند روش DWT در این روش تصویر در تجزیه موجک به چهار زیر تصویر تقسیم می‌شود. در این روش پس از منطبق کردن دو تصویر و باز نمونه‌برداری مجدد تصویر چند طیفی در ابعاد تصویر پانکروماتیک، باندهای تصویر چند طیفی تولید شده به فضای IHS می‌روند. در مرحله بعدی تطابق هیستوگرام تصویر پانکروماتیک و مؤلفه شدت در فضای IHS صورت می‌گیرد و تصویر پانکروماتیک جدید تولید می‌شود. سپس تجزیه موجک بر روی پارامتر شدت و تصویر پانکروماتیک جدید اعمال شده و ضرایب موجک جدیدی برای آن‌ها تولید می‌شود. در اینجا ضریب تقریب تصویر شدت تجزیه شده (LL<sub>۱</sub>) به جای ضریب تقریب تصویر پانکروماتیک تجزیه شده (LL<sub>Pan</sub>) قرار می‌گیرد. سپس تبدیل معکوس موجک به آن اعمال شده و نتیجه‌اش تولید تصویر جدیدی

---

<sup>۱</sup> HIS-Fast Fourier Transform

<sup>۲</sup> Grided Filter PCA

<sup>۳</sup> Ling

<sup>۴</sup> Zhang



از شدت می‌شود. درجات خاکستری این تصویر توزیع یکسانی با مؤلفه اصلی شدت دارد و شامل جزئیات مکانی یکسانی با تصویر پانکروماتیک است. در نهایت با یک تبدیل معکوس IHS تصویر تلفیق شده به دست می‌آید [۳۰].

روش GFPCA: در این روش از فیلترهای هدایت‌شده‌ای استفاده شده که با حفظ اعتدال بین توان تفکیک مکانی و طیفی یکی از چالش‌های اصلی در تلفیق تصاویر چند طیفی را برطرف می‌کند. این فیلترها علاوه بر حفظ اطلاعات طیفی از تصاویر چند طیفی، ساختارهای مکانی تصویر با وضوح مکانی بالا را به تصویر چند طیفی انتقال می‌دهند. برای افزایش سرعت پردازش این روش ابتدا از PCA برای تصویر چند طیفی استفاده شده و سپس محتوای اطلاعات از نویز جدا می‌شود [۳۱].

## ۲-۱-۴- روش‌های بیزی و فاکتورهای مبتنی بر ماتریس

دسته چهارم از روش‌های تلفیق روش‌های بیزی و فاکتورهای مبتنی بر ماتریس هستند که هدف اصلی روش‌های این دسته، تلفیق تصاویر چند طیفی و تصاویر پانکروماتیک در چارچوب استنتاج بیزی است. در این دسته تعیین مشکل توزیع مناسب هر کدام از عناصر موجود به عنوان یکراه حل مناسب برای حل مسئله معرفی شده است. یکی از روش‌های معروف این دسته روش فاکتور مبتنی بر تجزیه ماتریس (CMNF<sup>۱</sup>) است که این روش می‌تواند در دو گام اصلی انجام شود. گام اول تشکیل یک‌پایه برای زیر فضای سیگنال تصویر چند طیفی و گام دوم استفاده از یک تجزیه. روش فاکتور مبتنی بر تجزیه ماتریس، از هر دو منابع داده برای جمع‌آوری اجزای خالص و نقشه‌های فراوانی استفاده با توان تفکیک به‌طور مرتب استفاده می‌کند [۳۲].

---

<sup>۱</sup> Coupled Non-negative Matrix Factorization

## ۲-۲-۲- برخی از تحقیقات انجام شده مبتنی بر روش های تلفیق

در سال ۲۰۱۲ برایان آلن جانسون<sup>۱</sup> و همکاران به مقایسه روش های تلفیق IHS، Brovey و SFIM پرداختند. این تلفیق بر روش تصاویر سنجنده ژئوآی-۲<sup>۱</sup> انجام شد. نتایج این مقایسه حاکی از آن بود که الگوریتم هایی مثل IHS و Brovey با شفاف سازی و حفظ اطلاعات مکانی بیشتر از باند پانکروماتیک، منجر به تقسیم بندی های دقیق تر مکانی در نتایج می شوند، در صورتی که الگوریتم SFIM با به حداقل رساندن اعوجاجات طیفی تصویر چند طیفی وضوح پایین منجر به بهتر بودن اطلاعات طیفی نتایج می شود. همچنین ایشان رویکردی جدیدی بر اساس یافته های بیان شده ارائه نمودند با نام HIS-SFIM که اطلاعات مکانی و طیفی بهتری در نتایج نسبت به روش های قبلی ارائه نموده است. [۳۳]

علی دوست و همکاران در سال ۲۰۱۳ با استفاده از اطلاعات مکانی تصویر آیکونوس<sup>۲</sup> با توان تفکیک مکانی ۴ متر و اطلاعات طیفی تصویر هایپریون<sup>۴</sup> با توان تفکیک مکانی ۳۰ متری، تصویری با توان تفکیک مکانی ۴ متر و ۱۳۶ باند تهیه کردند. ایشان از تلفیق به روش جداسازی مکانی خطی استفاده نمودند و برای ارزیابی نتایج نیز، مقایسه ای بین پیکسل بازسازی شده تصویر هایپریون و پیکسل واقعی صورت گرفت [۵].

در سال ۲۰۱۶ محمد نژاد نیازی و همکاران روشی جدید بجای روش مشهور IHS<sup>۵</sup> باهدف کاهش اعوجاجات مکانی در تصویر شدت برای مناطق گیاهی پیشنهاد نمودند. در این روش از شاخص گیاهی تفاضلی نرمال شده جهت شناسایی مناطق پوشش گیاهی استفاده شده است. در طی این فرآیند باند سبز به کمک ترکیبی از باندهای قرمز و مادون قرمز بهبود داده شد. در روش معرفی شده در این تحقیق از الگوریتم بهینه سازی ژنتیک برای یافتن بهترین

---

<sup>۱</sup> Brian Alan Johnson

<sup>۲</sup> GeoEye-1

<sup>۳</sup> IKONOS

<sup>۴</sup> Hyperion

<sup>۵</sup> Intensity Hue Saturation

وزن تصویر شدت استفاده شد. آنالیزهای بصري و آماری توسط روش‌های ارزیابی کیفیت طیفی و مکانی حاکی از آن است که روش مورد بحث توانسته کیفیت تصویر تلفیق شده را نسبت به خروجی روش‌های مرسوم بهبود دهد. [۳۴].

در سال ۲۰۱۷ بشیر پور و همکاران مقایسه‌ای بین روش‌های تلفیق تصاویر ۳۰ متری هایپریون و تصاویر پانکروماتیک ۲/۵ متری کارتوست<sup>۱</sup>-۱ انجام دادند. این مقایسه بین روش‌های Wavelet-PCA[29]، FFT-PCA[29]، Wavelet-IHS و FFT-IHS صورت گرفته است. ارزیابی عددی نیز در معیارهای ضریب همبستگی ( $CC^2$ )، ضریب همبستگی فیلتر شده ( $FCC^3$ )، خطای کمترین مربعات ( $RMSE^4$ )، شاخص ERGAS و معیار نگاشت زاویه طیفی ( $SAM^5$ ) انجام شده است. با توجه به نتایج عددی، روش‌های مبتنی بر PCA دقت بالاتری در تلفیق تصاویر نسبت به IHS داشته‌اند. در حقیقت روش‌های مبتنی بر PCA به علت نداشتن محدودیت باندی دارای مزایای بیشتری نسبت به روش‌های IHS هستند (شکل (۲-۳)) [۲۶].

---

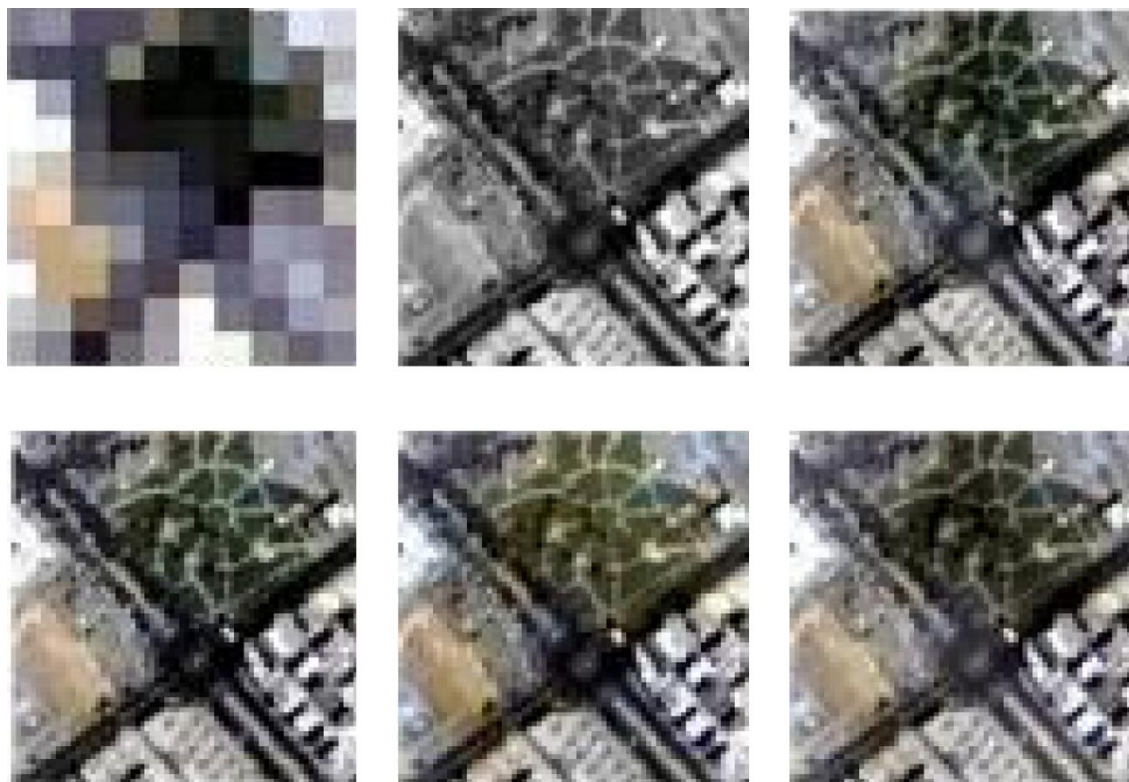
<sup>1</sup> Cartosat

<sup>2</sup> Correlation Coefficients

<sup>3</sup> Filter Correlation Coefficients

<sup>4</sup> Root-mean-square deviation

<sup>5</sup> Spectral angle mapping



شکل ۲-۳: تصاویر به ترتیب چپ به راست و بالا به پایین. تصویر چند طیفی هایپریون، تصویر پانکروماتیک کارتوست-۱، روش های تلفیق FFT-PCA، Wavelet-IHS، FWavelet-PCA و FFT-IHS

در سال ۲۰۱۸ گوردانا کاپلان<sup>۱</sup> و همکاران به مطالعه بر روی افزایش وضوح باند های تصویر سنجنده سنتینل-۲ پرداختند. سنجنده سنتینل-۲ فاقد باند پانکروماتیک است. به همین ترتیب ایشان سه روش تولید باند پانکروماتیک را مورد مقایسه قرار داد. در روش اول از نزدیک ترین باند با وضوح مکانی بالا به وضوح مکانی پایین، به عنوان باند پانکروماتیک استفاده شد. در روش دوم با میانگین گیری از تمامی باند های وضوح بالا یک باند پانکروماتیک تولید نمودند. در روش سوم از همبستگی خطی برای تعیین باند پانکروماتیک استفاده شد. در این تحقیق باند های ۶۰ متری مورد مطالعه قرار نگرفتند. با در دسترس بودن تصویر پانکروماتیک از سه روش تلفیق

<sup>۱</sup> Gordana Kaplan

تصاویر (IHS، HPF و WPC) برای مقایسه و ارزیابی عددی روش های معرفی شده استفاده شده است. نتایج نشان می دهد که استفاده از میانگین باند های وضوح بالا بهترین باند پانکروماتیک را تولید می کند. [۳۵]

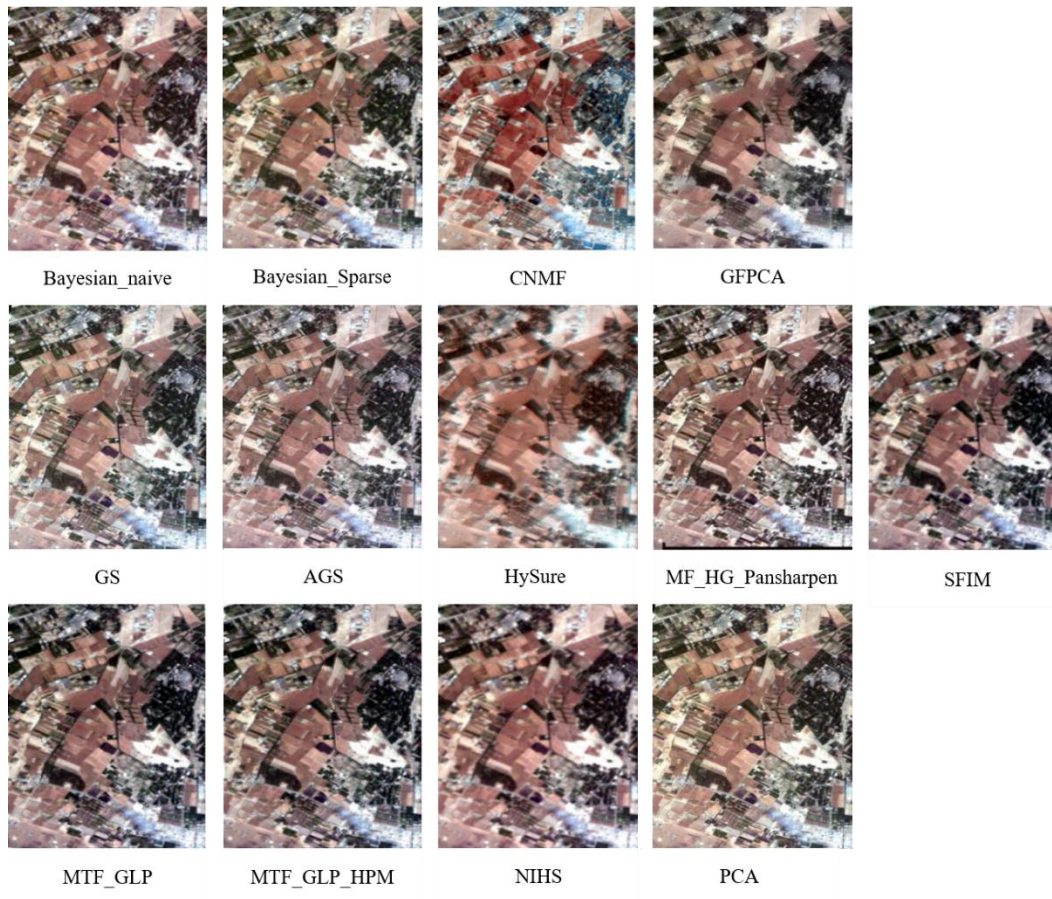
در سال ۲۰۱۹ شاه حسینی و همکاران به تحلیل مقایسه ای دقت و کارایی روش های نوین تلفیق داده در مناطق کشاورزی و شهری با استفاده از تصاویر فرا طیفی با انواع کلاس های پیچیده و متنوع به صورت بصری و کمی پرداختند. ایشان در روند تلفیق از تصویر چند طیفی هایپریون و تصویر پانکروماتیک مربوط به سنجنده (ALI<sup>۱</sup>) استفاده نمودند. نتایج نشان داد که روش مبتنی بر تجزیه ماتریس (CNMF<sup>۲</sup>) عملکرد بهتری در مقایسه با سایر روش ها دارد، به این گونه که توانسته توان تفکیک عوارض را در سطح پیکسل با حفظ جزئیات طیفی و مکانی بهبود بخشد. این اما این روش نیازمند زمان پردازش بسیار بالایی است. از سویی روش (NHIS<sup>۳</sup>) نیازمند زمان پردازش کمتری دارد ولی در حفظ جزئیات عملکرد مناسبی نداشته است (شکل (۲-۴)) [۲۰].

---

<sup>۱</sup> Advance Land Imager

<sup>۲</sup> Factorization Coupled Non-negative Matrix

<sup>۳</sup> Non-linear Intensity Hue-Saturation



شکل ۲-۴: نتایج بصری روش‌های مختلف تلفیق

در سال ۲۰۱۹ کابلی زاده و همکاران برای پایش شهر اهواز، از تلفیق تصاویر سنجش از دور لندست-۸ و سنتینل-۲ به کمک روش‌های تلفیق تصاویر، گرام اشمیت<sup>۱</sup>، CN<sup>۲</sup>، Brovey<sup>۳</sup>، DWT<sup>۳</sup>، SFIM<sup>۴</sup> و Multiplication پرداختند. نتایج نشان داد که روش‌های DWT و Brovey به ترتیب بهترین روش از لحاظ حفظ اطلاعات طیفی و مکانی بوده و

<sup>۱</sup> Gram Schmidt

<sup>۲</sup> Color Normalized

<sup>۳</sup> Discrete Wavelet Transform

<sup>۴</sup> Smoothing Filter-based Intensity Modulation

بیشترین اعوجاجات از نظر طیفی مربوط به روش Multiplication می باشد. ایشان نشان دادند که روش SFIM از جنبه طیفی و مکانی دارای دقت مناسب است [۳۶].

در سال ۲۰۲۱ خاویر رایموند<sup>۱</sup> و همکاران برای بهبود وضوح طیفی و مکان تصاویر حرارتی اخذ شده توسط هواپیمای بدون سرنشین (UAV<sup>۲</sup>) از الگوریتم های پنشارپنینگ تصاویر استفاده نموده و در ادامه نتایج آن را مورد تجزیه و بررسی قرار دادند. به کمک روش های کمی، مقایسه ای بین روش های پنشارپنینگ در تلفیق تصاویر با وضوح بالا و تصاویر حرارتی بدست آمده از پهپاد ها انجام شد تا بتوانند بهترین روش را جهت بهبود وضوح تصاویر حرارتی پهپاد ها پیدا کنند. روش های استفاده شده برای تلفیق تصاویر حرارتی پهپاد و تصاویر وضوح بالا عبارتند از: IHS<sup>۳</sup>، PCA<sup>۴</sup>، BDS<sup>۵</sup>، GS<sup>۶</sup>، PRACS<sup>۷</sup>، HPF<sup>۸</sup>، SFIM<sup>۹</sup>، INDUSION<sup>۱۰</sup>، MTF-GLP<sup>۱۱</sup>، MTF-GLP-HPM<sup>۱۲</sup>، MTF-GLP<sup>۱۳</sup>، MTF-GLP-ECB<sup>۱۴</sup> و HPM-PP<sup>۱۵</sup> که روش BDS بهترین نتایج را نسبت به روش های دیگر ارائه کرده است. [۳۷]

در سال ۲۰۲۱ یعقوبی و همکاران از فیلترگذاری نرم با ابعاد بهینه در فضای فرکانس به منظور بهبود تلفیق تصاویر چند طیفی و پانکروماتیک استفاده نمودند. در این راستا به منظور بهره مندی توأم از اطلاعات طیفی و جزئیات مکانی تصاویر سنجش از دوری، شاخصی وزن دار متشکل از دو جزء رنگ پذیری و جزئیات مکانی پیشنهاد نمودند که توانسته با تکیه بر معیارهای هندسی و آماری فضای ویژگی در کنار فیلترهای نرم، نتایج مطلوبی را در شناسایی ابعاد بهینه فیلترگذاری داشته باشد. در این سازوکار، وزن شاخص های رنگ پذیری و جزئیات اطلاعات مکانی با مدنظر

---

<sup>1</sup> Javier Raimundo

<sup>2</sup> Unmanned Aerial Vehicle

<sup>3</sup> Band-Dependent Spatial-Detail

<sup>4</sup> Partial Replacement Adaptive Component Substitution

<sup>5</sup> High-Pass Filtering

<sup>6</sup> Decimated Wavelet Transform (DWT) using an additive injection model

<sup>7</sup> Generalized Laplacian Pyramid with Modulation Transfer

Function

<sup>8</sup> GLP with MTF-matched filter and multiplicative injection

model

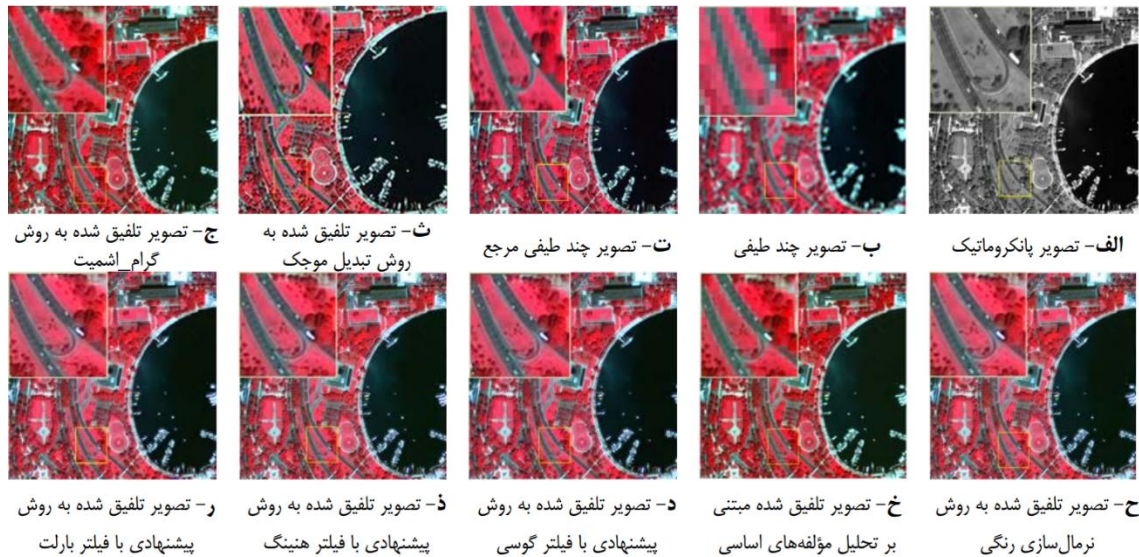
<sup>9</sup> GLP with MTF-matched filter, multiplicative injection model

and post-processing

<sup>10</sup> MTF-GLP with Enhanced Context-Based model

داشتن تفاوت در محتوای تصاویر، برای هر تصویر به صورت جداگانه برآورد می‌شود. نتایج عددی، برتری ۵۸ درصدی روش پیشنهادی (روش تبدیل موجک) را در رقابت با روش‌های دیگر در معیار RMSE نشان می‌دهد (شکل ۲-۵).

[۳۸].



شکل ۲-۵: نتایج تصاویر تلفیق شده روش‌های مختلف

## ۲-۲- تحقیقات پیشین مربوط به روش‌های سوپرزولوشن

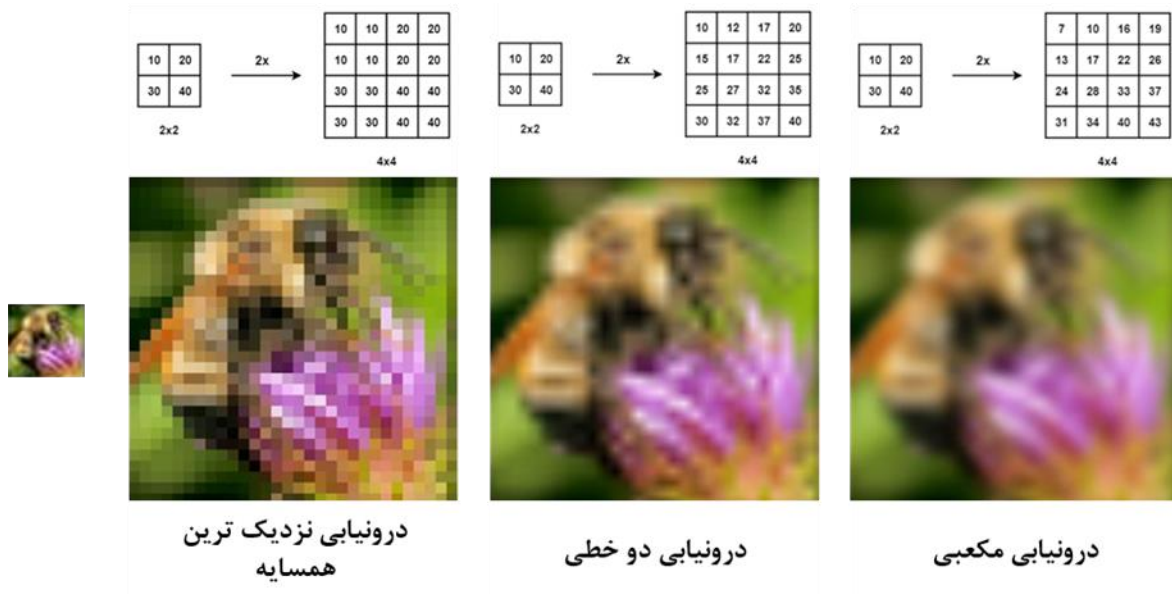
روش‌های سوپرزولوشن به طور کلی به دودسته تقسیم‌بندی می‌شوند. (۱) روش‌های مبتنی بر درون‌یابی، (۲) روش‌های مبتنی بر یادگیری. در ادامه به معرفی و بررسی روش‌های معرفی شده و جزئیات آن‌ها پرداخته شده است.

### ۲-۳-۱- روش‌های مبتنی بر درون‌یابی

یکی از قدیمی‌ترین روش‌ها در حوزه سوپرزولوشن، روش‌های مبتنی بر درون‌یابی هستند. در این روش‌ها عموماً از یک عملگر پایه درون‌یابی برای بازسازی نقاط خالی شبکه منظم مربوط به تصویر با وضوح بالا، استفاده می‌شود.



مهم‌ترین مزیت این روش‌ها سادگی و پیچیدگی محاسبات کم است. روش درونیابی نزدیک‌ترین همسایه<sup>۱</sup> که یک درونیابی مرتبه صفر است، ساده‌ترین و سریع‌ترین روش در میان این روش‌ها قلمداد می‌شود. در این روش هر پیکسل مجهول با مقدار موجود در نزدیک‌ترین همسایه‌اش پر می‌شود که باعث تخریب همواری تصویر و رخداد حالت شطرنجی (دندانه‌دندانه) در نتایج می‌گردد [۳۹]. برای حل این مسئله روش‌های درونیابی با مرتبه نرمی بالاتر پیشنهاد شده (مانند درونیابی دوخطی<sup>۲</sup> و مکعبی<sup>۳</sup>) که برای تخمین درجات خاکستری در تصاویر ارتقاء یافته از تعداد بیشتری همسایه از تصویر مبنا استفاده می‌نمایند (شکل (۶-۲)). در میان روش‌های درونیابی خطی ای که معرفی شده‌اند، درونیابی SPLINE بهترین نتایج را ارائه کرده است [۳۹].



شکل ۶-۲: نمایی از نحوه عملکرد روش‌های درونیابی

این موضوع کاهش اثر شطرنجی شدن را به دنبال داشته اما بروز ماتنی در محل لبه‌ها را به دنبال دارد [۴۰]. برای کاهش این چالش نیز روش‌های درونیابی جهتی توسعه یافته که در این روش‌ها ابتدا در هر نقطه یک جهت مشخص

<sup>۱</sup> Nearest Neighbor

<sup>۲</sup> Bilinear

<sup>۳</sup> Bicubic

می‌شود که در آن جهت تصویر تغییرات حداقلی دارد. سپس از نظم هندسی ساختارهای تصویر در این جهت خاص، برای درون‌یابی تطبیقی داده‌ها در هر نقطه استفاده می‌شود. به این صورت که در نواحی لبه‌ها، جهت درون‌یابی باید موازی با لبه‌ها باشد. به این ترتیب معمولاً این روش‌ها از آشکار ساختن لبه و عملگر گرادیان، برای مشخص نمودن جهت درون‌یابی استفاده می‌نمایند که باعث بهبود کیفیت لبه‌ها نسبت به روش‌های قبلی می‌شود [۶]. روش دیگری که در این زمینه معرفی شده است NEDI<sup>۱</sup> نام دارد که اولین بار در سال ۲۰۰۱ توسط لی<sup>۲</sup> معرفی شد. در این روش ابتدا ضرایب کوواریانس تصویر وضوح پایین (LR) تخمین زده می‌شوند و سپس به کمک دوگانگی هندسی بین کوواریانس تصویر LR و تصویر HR، پروسه درون‌یابی انجام می‌شود و تصویر HR به وجود می‌آید [۴۱].

## ۲-۳-۲- روش‌های مبتنی بر یادگیری

در این روش‌ها برای تولید تصویر HR از تصویر LR لازم است ابتدا شبکه موردنظر به کمک داده‌های بسیار زیادی آموزش داده شود. این آموزش می‌تواند داخلی یا خارجی باشد. به عبارت بهتر آموزش شبکه می‌تواند به کمک داده‌های منطقه مورد مطالعه یا مناطق دیگر صورت پذیرد که بر همین اساس شبکه‌های متعددی در سال‌های اخیر معرفی شده است. در ادامه ابتدا مفاهیمی پایه از شبکه‌های عصبی تبیین شده و سپس برخی از تحقیقات انجام شده در این زمینه مورد بررسی قرار گرفته است.

## ۲-۳-۲-۱- مروری کوتاه بر مفاهیم پایه روش‌های سوپرزولوشن

در تولید تصویر HR از تصویر LR به روش‌های مبتنی بر یادگیری لازم است الگوریتمی ارائه شود که توانایی یادگیری وضوح بالاتر را با استفاده از تعداد مشخصی داده آموزشی داشته باشد. به‌طور کلی روش‌های یادگیری ماشین می‌توانند در حل مسائل مختلف با بهینه‌سازی پارامترها در روند آموزش بسیار کارآمد باشند [۴۲]. در بین

---

<sup>۱</sup> New Edge Directed Interpolation

<sup>۲</sup> Li

الگوریتم‌های یادگیری ماشین<sup>۱</sup>، شبکه‌های عصبی مصنوعی<sup>۲</sup> که بر پایه ساختار بیولوژیکی عصبی ذهن انسان طراحی شده‌اند و از ساختاری نظارت‌شده<sup>۳</sup> تشکیل شده‌اند، می‌توانند راه‌حل مناسبی جهت حل مسئله موردنظر در این تحقیق باشند. اما از آنجاکه یافتن رابطه بین تصویر LR و تصویر HR بسیار پیچیده بوده و از لحاظ محاسباتی مسئله‌ای بد وضع است، حل آن نیازمند به یک ساختار عمیق با توانایی یادگیری الگوهای متفاوت موجود در تصویر می‌باشد [۴۲]. یادگیری عمیق شاخه‌ای از یادگیری ماشین است که از ساختار آموزش سلسله مراتبی بوده تشکیل شده و پارامترها در سطوح پایین‌تر به‌منظور توصیف پارامترها در سطوح بالاتر استفاده می‌شوند [۴۳]. امروزه از یادگیری عمیق در حوزه ماشین بینایی به‌صورت گسترده مورد بحث و مطالعه قرار گرفته و در زمینه‌های مختلف و کاربردهای متنوع همچون، شناسایی اشیاء [۴۴]، پردازش زبان طبیعی [۴۵]، طبقه‌بندی تصاویر [۴۵]، برچسب‌گذاری معنایی [۴۶] و تشخیص صدا [۴۷] مورد استفاده قرار گرفته است. بنابراین، استفاده از شبکه‌های عمیق به‌عنوان روشی که توانایی‌های لازم جهت استفاده در زمینه افزایش وضوح تصاویر سنجش‌ازدوری به‌عنوان گزینه‌های حل مسئله مطرح است. این شبکه‌ها هم توانایی استخراج ویژگی‌های مختلف در سطوح متفاوت را داشته و هم توانایی یادگیری با استفاده از تعدادی داده آموزشی را دارند [۴۸]. در ادامه این بخش به‌مرور مفاهیم پایه در شبکه‌های عصبی مصنوعی و شبکه‌های عمیق پرداخته شد است.

## ۲-۳-۱-۱-۱- شبکه‌های عصبی مصنوعی

این شبکه‌ها به‌عنوان نسل جدید دستگاه‌های پردازش اطلاعات بسیار مورد توجه قرار گرفته‌اند. یک شبکه عصبی توانایی‌هایی نظیر یادگیری، تعمیم، برطرف نمودن نویز و پردازش موازی می‌باشد. در ساختار، شبکه عصبی دارای یک لایه ورودی، یک یا چند لایه میانی و یک لایه خروجی می‌باشد. تنوع شبکه‌های عصبی با توجه به ارتباط میان عصب‌ها، الگوریتم یادگیری شبکه و تابع محرک تعیین می‌شوند. شبکه‌های عصبی دارای معماری‌های متفاوت

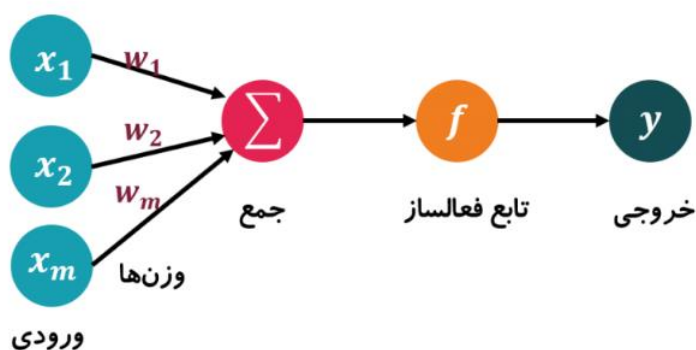
---

<sup>۱</sup> Machine learning

<sup>۲</sup> Artificial neural network

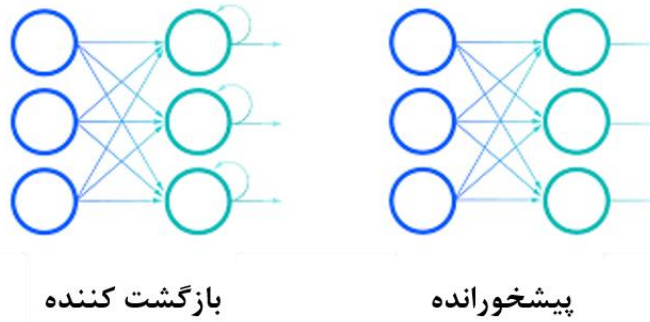
<sup>۳</sup> Supervised

می باشند و به طور کلی به دو گروه پیش خورنده<sup>۱</sup> و بازگشت کننده تقسیم می شوند [۴۸]. شبکه های پیش خورنده صرفاً محاسبات روبه جلو خواهد بود و بازگشت ندارد، اما در شبکه های بازگشت کننده، خروجی هر لایه می تواند به لایه قبلی خودش و یا لایه بعدی راه پیدا کند (شکل (۷-۲))



شکل ۷-۲: ساختار شبکه عصبی

شکل (۶-۲) ساختار کلی شبکه های عصبی را به نمایش گذاشته است در لایه اول این ساختار داده های ورودی قرار دارند که در وزن مربوط به خودشان ضرب شده و در قسمت گره باهم جمع می شوند. نتیجه وارد یک تابع نیمه خطی می شود و در نهایت خروجی حاصل می شود. شبکه پیش خورنده ساده ترین نوع شبکه عصبی است، زیرا خروجی هر لایه فقط وابسته به ورودی لایه می باشد. در شبکه های عصبی بازگشت کننده خروجی در هر لحظه علاوه بر سیگنال های ورودی در همان لحظه و وزن های عصبی، به خروجی یا برخی از خروجی های لحظات ماقبل نیز وابسته است [۴۸] (شکل (۸-۲)).



شکل ۸-۲: مقایسه شماتیک شبکه عصبی پیش خورانده و بازگشت کننده.

هر چه تعداد لایه‌های شبکه عصبی بیشتر باشد، انعطاف‌پذیری شبکه در مسائل پیچیده بیشتر می‌شود و به طبع محاسبات پیچیده و وقت‌گیرتر می‌شود [۴۸]. پروسه آموزش شبکه عصبی به این صورت است که در هر تکرار با به‌روزرسانی وزن‌ها و بایاس‌ها مقادیر بهینه این پارامترها برای رسیدن به بهترین نتایج به دست می‌آید. شکل (۲-۹) یک کد مربوط به شبکه‌های عصبی می‌باشد [۴۲].

1. Input: Input  $x$ , Loss function  $L$ , Ground truth  $y$ , Epochs  $K$ , Weights  $W$ ,
2. Output: Network prediction  $\hat{y}$ ,
3. Dataset  $\Omega$ : Training set  $\Omega_{train} \subset \Omega$ , Validation set  $\Omega_{val} \subset \Omega$ , Test set  $\Omega_{test} \subset \Omega$  with  $\Omega_{test} \cap \Omega_{val} \cap \Omega_{test} = \emptyset$ .
4. Initialization: Random weights, epochs count  $i = 0$ .
5. while  $i \leq k$  do
6. Forward propagate  $x \in \Omega_{train}$  through network
7. Generate prediction  $y$  at output
8. Evaluate prediction  $L(y, \hat{y})$
9. Backpropagation of loss gradient  $\frac{\partial L(y, \hat{y})}{\partial w}$
10. Update  $W$  and  $b$  with gradient based optimization method
11. Evaluate network on  $x \in \Omega_{val}$
12. end
13. Generate predictions of test set  $x \in \Omega_{test}$  with the network with best performance on validation set
14. Evaluate performance of network predictions

شکل ۹-۲: شبه کد یک شبکه عصبی

مقادیر وزن‌ها و بایاس‌ها در شبکه‌های عصبی طبق روابط (۲-۴ و ۲-۵) به‌روزرسانی می‌شوند.

$$W_{i,j}^m(k+1) = W_{i,j}^m(k) + 2\alpha(t_i - \alpha_i)f^m(n_i^n)W_j^{m-1} \quad \text{رابطه (۲-۴)}$$

$$b_i^m(k+1) = b_i^m(k) + 2\alpha(t_i - \alpha_i)f^m(n_i^n) \quad \text{رابطه (۲-۵)}$$

در این روابط معرفی شده  $b$  بایاس،  $w$  وزن،  $m$  شماره لایه،  $\alpha$  نرخ یادگیری<sup>۱</sup>،  $k$  تعداد تکرار،  $t_i$  نتیجه پیش فرض متناظر با ورودی هر لایه،  $\alpha_i$  خروجی لایه  $i$  ام،  $f$  تابع فعال‌ساز مربوط به هر لایه و  $n$  متغییر وابسته به  $f$  می‌باشد [۴۸].  
روش گرادیان کاهشی<sup>۲</sup> روشی است جهت بهینه‌سازی سلسله مراتبی که بر مبنای گرادیان می‌باشد و در راستای کمینه کردن تابع اتلاف مورد استفاده قرار می‌گیرد. بنابراین با افزایش تدریجی گام‌های پردازش در نقطه مورد نظر در جهت منفی گرادیان و حرکت به سمت کمینه تابع، بهینه‌سازی صورت می‌گیرد (رابطه (۲-۶)) [۴۸].

$$x_{i+1} = x_i - \gamma \nabla F(x_i) \quad \text{رابطه (۲-۶)}$$

در رابطه (۲-۶)  $x_i$  مقدار به‌دست‌آمده نقطه در محل محاسبه تابع اتلاف و  $\gamma$  بزرگی گام‌های پارامتر است.  
مرحله بعدی پس انتشار<sup>۳</sup> بوده که یکی از مراحل کلیدی و مهم در گسترش شبکه‌های عصبی است. پس انتشار باعث آسان شدن محاسبه گرادیان محلی می‌شود. در این پروسه برای تجزیه گرادیان تابع اتلاف از قانون زنجیره‌ای استفاده می‌شود. به‌طور کلی در ابتدا، شبکه عصبی به کمک ساختار پیشروند مقداری تخمین زده می‌شود. سپس تابع اتلاف از مقایسه مقدار تخمین زده شده با مقدار مرجع به دست می‌آید. سپس تابع اتلاف با توجه به مقادیر وزن‌های محاسبه‌شده و مقدار خروجی تولیدشده، توسط پروسه پس انتشار به سمت اول شبکه بازمی‌گردد. این روند تا جایی ادامه پیدا می‌کند که مقادیر بهینه و خطاها کمینه شوند [۴۸].

---

<sup>۱</sup> Learning Rate

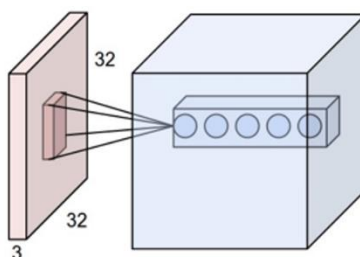
<sup>۲</sup> Gradient descent

<sup>۳</sup> Back propagation

## ۲-۳-۱-۲- لایه‌های کانولوشنی

یکی از مهم‌ترین بخش‌های تشکیل‌دهنده شبکه عصبی لایه‌های کانولوشنی هستند که خروجی این لایه‌ها به صورت یک توده سه بعدی دارای طول، عرض و عمق می‌باشند. انجام کانولوشن سه مزیت مهم دارد: (۱) مکانیزم اشتراک وزن در هر نقشه ویژگی باعث کاهش شدید تعداد پارامترها می‌گردد، (۲) اتصال محلی سبب آموزش ارتباط بین پیکسل‌های همسایه می‌گردد و (۳) این لایه باعث تغییر ناپذیری و ثبات نسبت به تغییر مکان عارضه می‌شود. به علت مزایای مطرح شده، در برخی شبکه‌ها از این لایه‌ها به عنوان جایگزین لایه‌های تماماً متصل استفاده نموده که باعث افزایش سرعت آموزش شبکه شده است. لایه‌های کانولوشنی انواع مختلفی دارند که در ادامه به بررسی برخی از آن‌ها پرداخته شده است [۴۹].

لایه کانولوشن ساده: در این مدل لایه‌ها هر لایه تشکیل شده از فیلترهایی است که حاوی وزن‌ها و بایاس می‌باشند. عملگرها ویژگی‌ها و اطلاعات بسیاری از ورودی لایه استخراج می‌کنند. که هر عملگر در قالب یک ماتریس با ابعاد  $F \times F$  و گام‌های  $S$  که این ابعاد و گام‌ها مشخص کننده تعداد وزن‌ها و ناحیه پذیرش در فیلتر می‌باشند. با ضرب این فیلترها به ورودی لایه نقشه ویژگی تولید می‌شود (شکل (۲-۱۰)).



شکل ۲-۱۰: نحوه اتصال در لایه کانولوشن

در شکل (۲-۹) ورودی تصویری با اندازه  $3 \times 32 \times 32$  و در سمت راست توده‌ای از نورون‌ها مشخص شده‌اند که هر نورون در لایه کانولوشن فقط به یک ناحیه از تصویر ورودی متصل است و به صورت عمقی باندهای مختلف را در

بردارد. با توجه به پارامترهای عمق، گام و پدینگ<sup>۱</sup> با صفر ابعاد خروجی مشخص می‌شود [۵۰]. ارتباط بین گام، نقشه ویژگی و ابعاد عملگر در رابطه (۷-۲) آورده شده است.

$$\begin{cases} X_{conv} = (X - F)/S + 1 \\ Y_{conv} = (Y - F)/S + 1 \end{cases} \quad \text{رابطه (۷-۲)}$$

در رابطه (۷-۲)  $F$  ابعاد عملگر،  $S$  اندازه گام و  $X$  و  $Y$  ابعاد ورودی می‌باشند. از طرفی بایاس موجود در تصویر به تعداد لایه‌های تولیدشده در نقشه ویژگی وابسته می‌باشد. رابطه مربوطه به محاسبه تعداد پارامترهای موجود در شبکه است.

$$N_d = (F \times F \times D) \times C + C \quad \text{رابطه (۸-۲)}$$

که در این رابطه  $C$  تعداد کانال موجود در هر لایه و  $N_d$  تعداد پارامترهای موجود در شبکه است. هنگامی که از لایه کانولوشنی استفاده می‌شود، نقشه ویژگی شده از نظر ابعادی کوچک‌تر از تصویر اولیه خواهد بود. برای تولید تصویری هم‌اندازه با تصویر اولیه نیاز است از پدینگ با صفر استفاده شود [۵۰]. در این پروسه ابعاد تصویر ورودی با توجه به ابعاد فیلتر بزرگ‌تر خواهد شد و این بزرگ‌تر شدن به کمک اضافه کردن سطرها و ستون‌های حاوی صفر در حاشیه تصویر میسر می‌شود. ابعاد نقشه ویژگی تولیدشده به کمک رابطه (۹-۲) قابل محاسبه است.

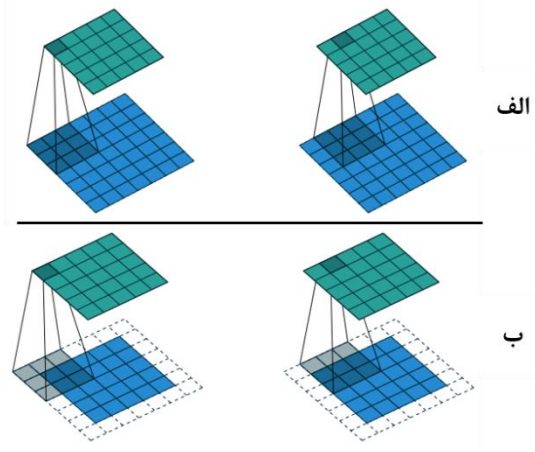
$$\begin{cases} X_{conv} = (X - F + 2P)/S + 1 \\ Y_{conv} = (Y - F + 2P)/S + 1 \end{cases} \quad \text{رابطه (۹-۲)}$$

در این رابطه مقدار  $P$  تعداد سطر و ستون در نظر گرفته‌شده برای پدینگ است. شکل (۱۱-۲) نحوه عملکرد کانولوشن ساده به کمک پدینگ با صفر و بدون استفاده از این روش نشان داده شده است.

---

<sup>۱</sup> Zero padding



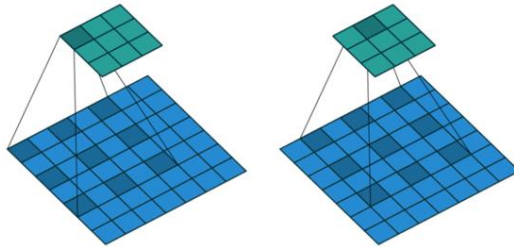


شکل ۱۱-۲: نحوه اثرگذاری پدینگ بر ابعاد نقشه ویژگی، (الف) بدون پدینگ، (ب) با پدینگ.

لایه کانولوشن متورم: این نوع از لایه کانولوشن برای افزایش ناحیه بدون افزایش پارامتر وزن معرفی شد [۸۳]. با فرض اینک  $F: Z^2 \rightarrow R$  یک تابع گسسته باشد،  $\Omega_r = [-r, r]^2 \cap Z^2$  و  $K: \Omega_r \rightarrow R$  عملگرهای گسسته با ابعاد  $(2r+1)^2$  باشند، رابطه (۱۰-۲) برای کانولوشن متورم به این صورت زیر تعریف می‌شود [۵۱].

$$(F *_l K)(q) = \sum_{s+lt} F(s)k(t) \quad \text{رابطه (۱۰-۲)}$$

که  $*$  نمایانگر عملگر کانولوشن متورم بوده | بیانگر نرخ تورم در کانولوشن متورم است (شکل (۲-۱۲)) [۵۱].



شکل ۱۲-۲: نحوه اعمال کانولوشن متورم با نرخ تورم ۲

لازم به ذکر است که در این کانولوشن نیز در صورت نیاز می‌تواند پروسه پدینگ اعمال شود.

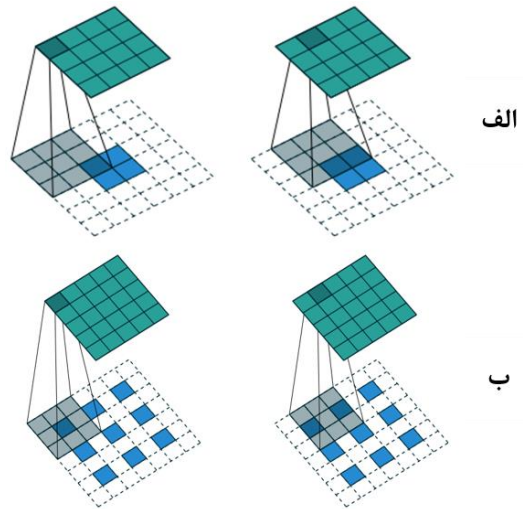
لایه کانولوشن معکوس<sup>۱</sup>: این نوع کانولوشن همانند نامش به طور معکوس عملکرد کانولوشن های قبلی عمل می کند. به این صورت که به جای کاهش ابعاد نقشه ویژگی باعث افزایش ابعاد ویژگی بدون استفاده از درون یابی می شود [۵۲]. به علت پیچیده بودن ارتباط میان لایه ها در شبکه عمیق کانولوشنی تماماً متصل، عمل کانولوشن معکوس بسیار دشوار خواهد بود (رابطه (۲-۱۱)، (۲-۱۲) و (۲-۱۳)).

$$E_{2 \times 2} = A_{3 \times 3} \times G_{4 \times 4} \quad \text{رابطه (۲-۱۱)}$$

$$E'_{4 \times 1} = A'_{4 \times 16} \times G'_{16 \times 1} \quad \text{رابطه (۲-۱۲)}$$

$$A'^T G' = A'^T A' E' \rightarrow G' = A'^T E' \quad \text{رابطه (۲-۱۳)}$$

در روابط بالا فرض بر این است که A ماتریس ورودی باشد، E نتیجه اعمال کانولوشن بر A و G عملگر کانولوشن است. حال اگر هدف برآورد ماتریس A باشد (فرضاً ابعاد فیلتر  $4 \times 4$ )، آنگاه ماتریس ورودی باید به فرم A' باشد. در این صورت ماتریس G به صورت برداری با ابعاد  $1 \times 16$  در ماتریس A ضرب خواهد شد و نتیجه یک بردار  $4 \times 1$  خواهد بود که به صورت  $2 \times 2$  نمایان می شود. لازم به ذکر است در کانولوشن معکوس نیست پارامترهای پدینگ، ابعاد فیلتر و گام در راستای تعیین ابعاد نقشه ویژگی مؤثر و قابل دست کاری خواهند بود (شکل (۲-۱۳)) [۴۲].



شکل ۲-۱۳: عملکرد کانولوشن معکوس، (الف) یک ماتریس  $2 \times 2$  با  $S=1$  و  $P=0$ ،  $F=3$ ، (ب) یک ماتریس  $3 \times 3$  با  $S=1$ ،  $P=0$ ،  $F=3$  و  $H=1$

### ۲-۳-۱-۲-۳- لایه ادغام<sup>۱</sup>

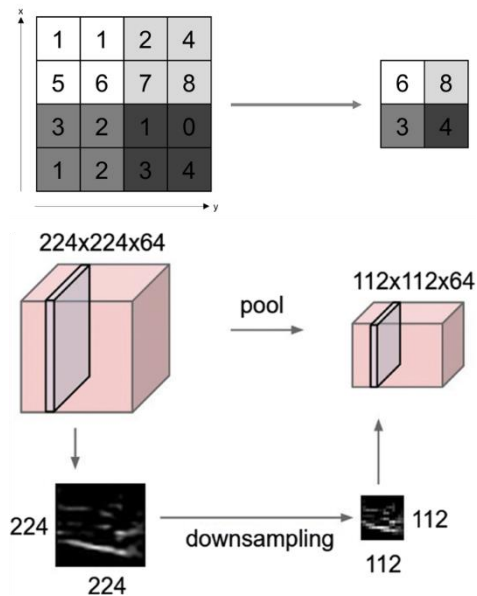
لایه بعدی بعد از لایه‌های کانولوشن معمولاً لایه ادغام است که عمدتاً جهت تغییر ابعاد نقشه ویژگی مورد استفاده قرار می‌گیرد. این لایه به دودسته لایه ادغام و لایه ادغام معکوس تقسیم‌بندی می‌شود [۴۲].

لایه ادغام: در این لایه از عملگرهای ماکزیمم یا میانگین استفاده می‌شود که هر کدام مزایای خود را داشته و با توجه به بررسی‌های صورت گرفته استفاده از توابع ماکزیمم باعث تسریع همگرایی و تعمیمی پذیری بیشتر می‌شود. در لایه ادغام ماکزیمم تعداد پارامترهای مکانی برای لایه‌های متوالی کاهش می‌یابد باعث افزایش ناحیه پذیرش می‌شود. این لایه از یک پنجره استفاده می‌کند که تصویر را به‌طور کامل اسکن می‌نماید [۴۲]. در فرایند هر اسکن مقدار ماکزیمم را ذخیره‌سازی می‌کند. اگر فرض شود ورودی این لایه  $X \times Y$  باشد ابعاد خروجی به شرح زیر خواهد بود (رابطه (۲-۱۴)).

<sup>۱</sup> Pooling

$$\begin{cases} X_{pool} = (X - F)/S + 1 \\ Y_{pool} = (Y - F)/S + 1 \end{cases} \quad \text{رابطه (۱۴-۲)}$$

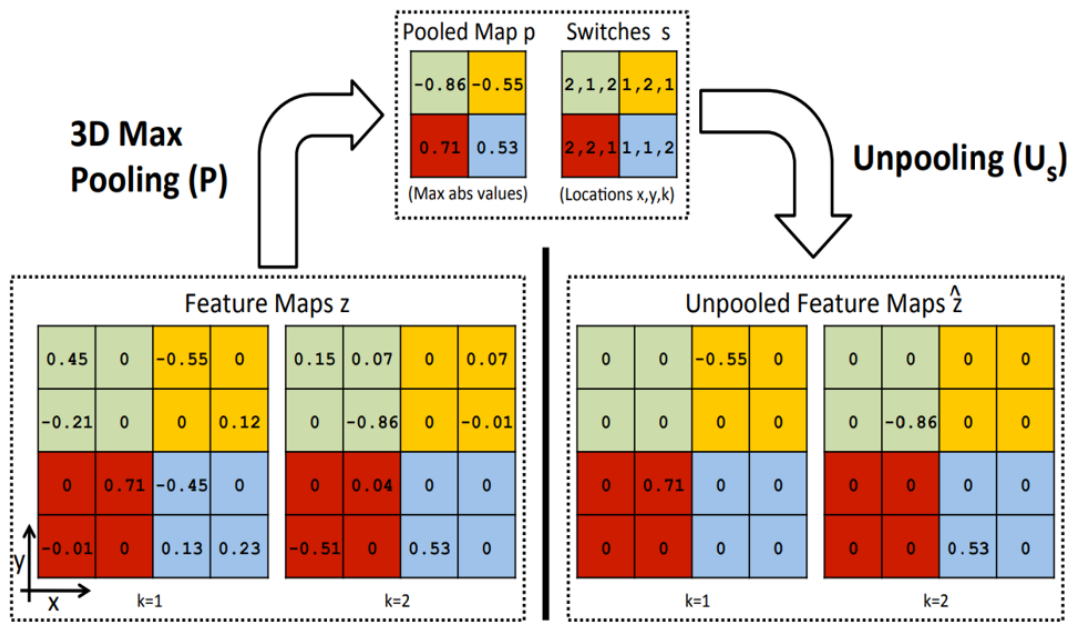
از طرفی عمق لایه‌های ورودی و خروجی با هم برابر خواهد بود. در مواردی که ابعاد موجود مناسب نمی‌باشند می‌توان از پدینگ استفاده نمود. رایج‌ترین ابعاد عملگرهای این لایه  $2 \times 2$  با گام‌های ۲ تایی می‌باشند (شکل ۲-۱۴).



شکل ۲-۱۴: عملکرد لایه ادغام ماکزیمم

لایه ادغام معکوس<sup>۱</sup>: این لایه کاملاً عکس لایه ادغام عمل می‌کند، یعنی باعث افزایش ابعاد ماتریس خروجی می‌شود. این پروسه باعث افزایش ابعاد ماتریس و حفظ خصوصیات کلی آن می‌شود [۴۲]. عملکرد این لایه در شکل (۲-۱۵) آورده شده است.

<sup>۱</sup> Unpooling



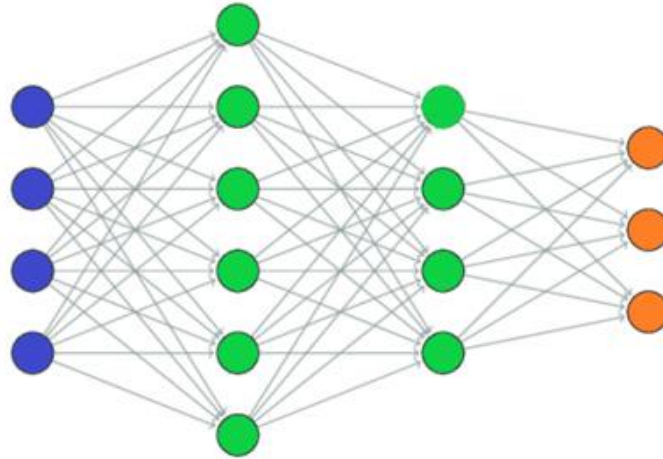
شکل ۱۵-۲: عملکرد لایه ادغام معکوس

در شکل (۲-۱۴) یک ماتریس  $4 \times 4$  با دولایه در عمق به نمایش گذاشته شده است که به کمک لایه ادغام ماکزیمم کوچک شده است و مقادیر موقعیت‌های آن ذخیره شده است. در لایه ادغام معکوس مقادیر ورودی در موقعیت‌های ذخیره شده جاگذاری شده و به باقی جایگاه‌ها صفر تعلق گرفته است.

### ۲-۳-۱-۴- لایه تمام‌متصل<sup>۱</sup>

لایه‌های تمام‌متصل آخرین لایه یک شبکه عصبی هستند که نقش آن‌ها تبدیل نقشه ویژگی به بردار است. به عبارت بهتر نتیجه شبکه پس از لایه‌های تمام‌متصل به صورت یک بردار با اندازه مشخص خواهد بود. این لایه‌ها دارای تعداد بسیار زیادی پارامتر هستند که باعث می‌شود هزینه پردازش و زمان پردازش زیاد شود و این امر یک ضعف بزرگ این لایه‌ها به شمار می‌رود (شکل (۲-۱۶)) [۵۳].

<sup>۱</sup> Fully connected



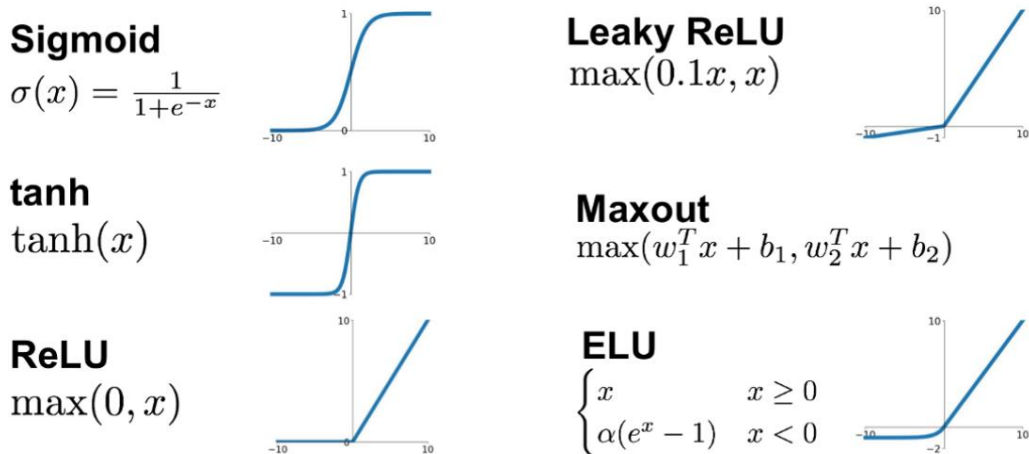
شکل ۱۶-۲: نحوه عملکرد لایه تمام‌متصل

## ۲-۳-۱-۵- توابع فعال‌سازی<sup>۱</sup>

این توابع جهت تبدیل خروجی در یک لایه از فرم خطی به غیر خطی استفاده می‌شوند. در بین توابع غیر خطی، تابع ReLU بیشترین محبوبیت را دارد. البته، از خانواده ReLU اعضای دیگری مانند PReLU، Leaky-ReLU و غیره وجود دارند. در شکل (۲-۱۷)، برخی از توابع غیر خطی قابل مشاهده است. برای مثال، تابع ReLU مقادیر منفی را صفر و مقادیر مثبت را بدون تغییر به خروجی انتقال می‌دهد [۵۴].

---

<sup>۱</sup> Activation function



شکل ۲-۱۷: نمونه‌هایی از توابع فعال‌ساز

## ۲-۳-۱-۶- مولفه‌های موثر در بهبود شبکه

مؤلفه‌هایی معرفی شده‌اند که در بهبود عملکرد شبکه عصبی در مسائل مختلف کمک می‌کنند. این مؤلفه‌ها عبارت‌اند از، ۱) لایه حذف تصادفی<sup>۱</sup>، ۲) لایه نرمال‌سازی دسته‌ای<sup>۲</sup>، ۳) انتقال آموزش<sup>۳</sup> و ۴) تقویت داده.

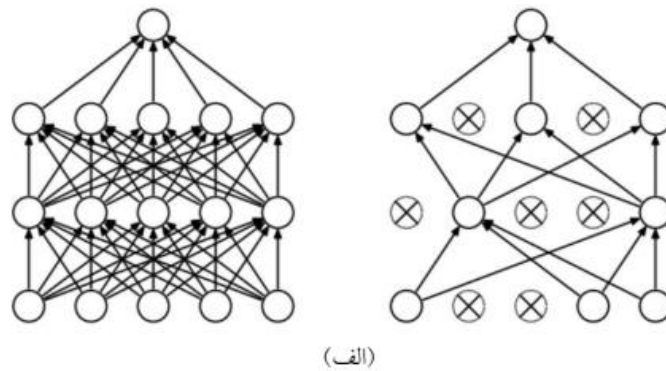
لایه حذف تصادفی: این لایه در راستای پیش‌گیری از بیش‌برازش<sup>۴</sup> شبکه مورد استفاده قرار می‌گیرد. در هر مرحله از آموزش، برای نرون‌ها است دو رویکرد با احتمالات  $p$  (جهت نگاه‌داشتن در شبکه) و  $1-p$  (جهت خروج از شبکه) موجود است که در نهایت موجب تولید یک شبکه کاهش داده شده می‌شود [۵۵]. همچنین اتصالات گره خارج‌شده نیز حذف می‌شود. به این صورت بخش باقی‌مانده در شبکه به کمک داده‌ها آموزش خواهد دید. سپس گره‌های حذف‌شده به همراه وزن‌هایشان مجدداً به ساختار شبکه بازخواهند گشت. این پروسه باعث تقویت آموزش شبکه می‌شود و شبکه را وادار می‌کند تا به کمک ویژگی‌های قدرتمند آموزش ببیند (شکل ۲-۱۸).

<sup>۱</sup> Dropout

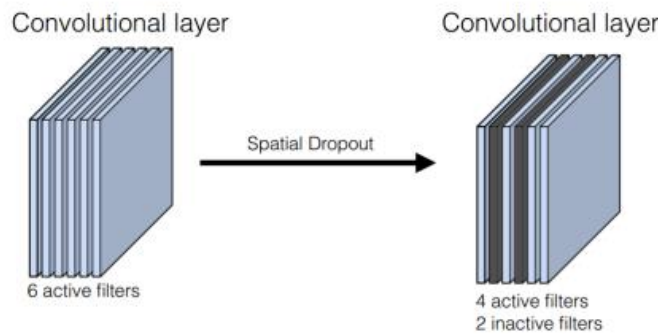
<sup>۲</sup> Batch Normalization

<sup>۳</sup> Transfer learning

<sup>۴</sup> Overfitting



(الف)



(ب)

شکل ۲-۱۸: (الف) پروسه حذف تصادفی در شبکه‌های معمولی، (ب) پروسه حذف تصادفی در شبکه‌های کانولوشنی.

لایه نرمال‌سازی دسته‌ای: یکی از پرکاربردترین لایه‌هایی که برای بهبود عملکرد شبکه مورد استفاده قرار می‌گیرند لایه نرمال‌سازی دسته‌ای می‌باشد. این لایه عمدتاً در مواقعی مورد استفاده است که مقادیر ویژگی‌های تولیدشده در بازه یکسانی نیستند. در این زمان به کمک این لایه تمامی مقادیر را به بازه یکسانی (مثلاً ۰ تا ۱ یا ۱- تا ۱) می‌برند تا زمان پردازش کوتاه‌تر و آسان‌تر شود [۵۶].

انتقال آموزش: شبکه‌های عصبی ذاتاً نیازمند داده‌های آموزشی بسیاری جهت یادگیری می‌باشند که تهیه داده آموزشی کافی همواره میسر نیست. در این مواقع استفاده از وزن‌های یک شبکه از پیش آموزش دیده به کمک مجموعه داده بزرگ یکراه حل کارساز و مؤثر می‌باشد. در عمل این وزن‌ها را به عنوان وزن اولیه وارد شبکه می‌نمایند و سپس با آموزش شبکه به کمک وزن‌های اولیه موجود قدرت تصمیم‌گیری شبکه را بهبود می‌دهند. [۵۶].



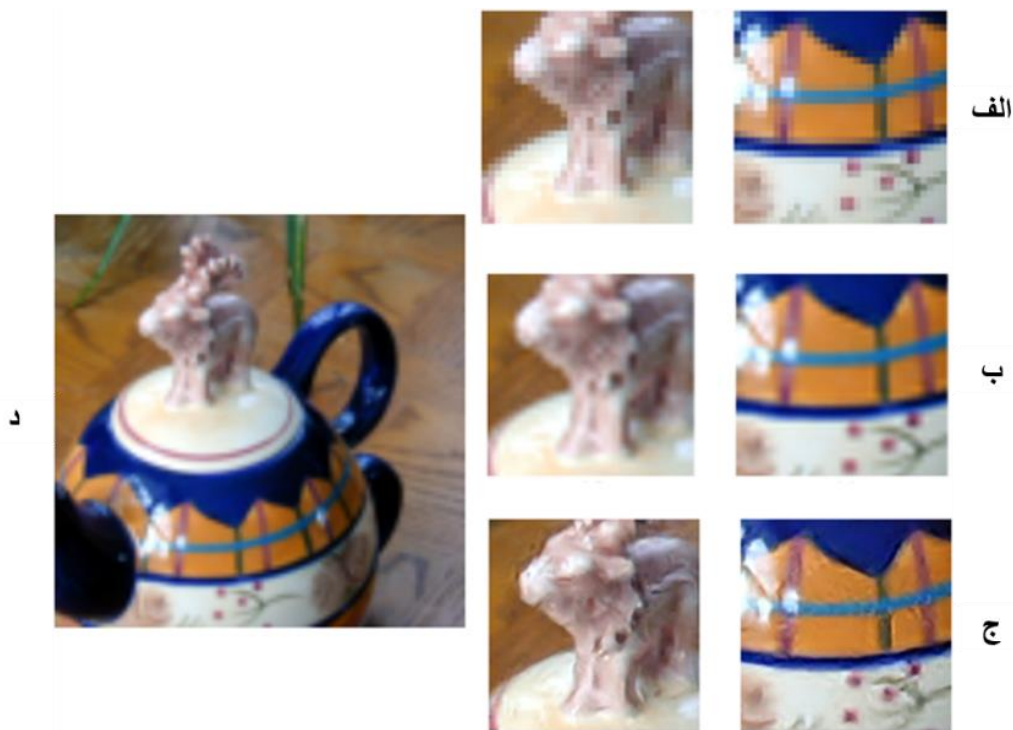
تقویت داده: تقویت داده یک رویکرد مؤثر در افزایش داده‌های آموزشی جهت افزایش دقت آموزش شبکه می‌باشد. این راهکار در زمانی مورد استفاده قرار می‌گیرد که داده‌های آموزشی محدود بوده و شبکه نیازمند تزریق داده آموزشی بیشتری می‌باشد. به‌طور کلی روش‌های تقویت داده به‌عنوان یک تنظیم‌کننده عمل می‌کنند و پروسه همگرایی را سریع‌تر می‌نمایند و از طرفی باعث بیش‌برازش شبکه نمی‌شوند. به‌طور کلی این رویکرد شامل تبدیلاتی همچون انتقال، چرخش، مقیاس‌گذاری، تغییر فضای رنگی، برش و غیره بر روی داده‌های آموزشی می‌باشد [۵۷]

## ۲-۳-۲-۲- برخی از تحقیقات انجام‌شده مبتنی بر روش‌های سوپررزولوشن

در سال ۲۰۰۲ فریمن<sup>۱</sup> و همکاران یک شبکه ملرکفی را برای تولید تصویر با وضوح مکانی بیشتر آموزش دادند. برای آموزش این شبکه از یک زوج داده شامل تصاویر HR و LR در موقعیت‌های جغرافیایی مختلف استفاده شده است. در این روش از تصویر لبه‌های استخراج‌شده از تصاویر LR نیز به‌عنوان لایه‌های ورودی آموزش شبکه استفاده شده. نتایج نشان می‌دهد این روش جزئیات ظریف مانند لبه‌ها را حفظ می‌کند و بافت‌های قابل‌باوری ایجاد نموده که می‌توان نتایج خوبی را پس از زوم کردن چندین اکتاو مشاهده نمود. نتایج مقایسه بصری برخی از نتایج شبکه ایشان با شبکه‌های رقیب در شکل (۲-۱۹) قابل‌مشاهده است [۵۸].

---

<sup>۱</sup> Freeman

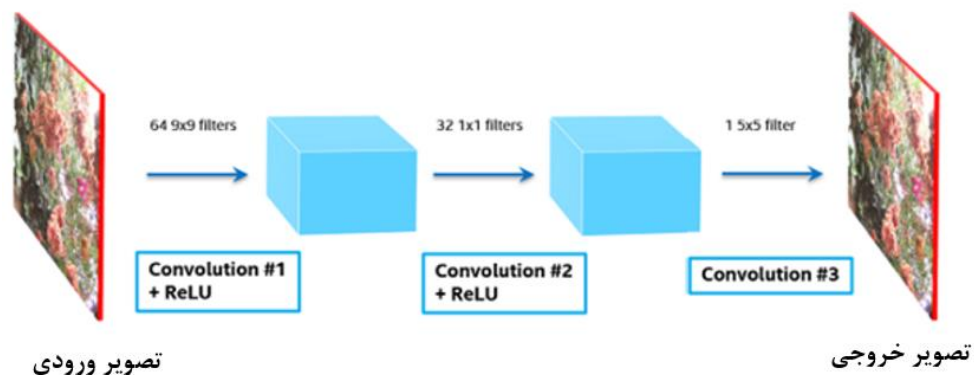


شکل ۲-۱۹: تصویر (د) تصویر اصلی، سطر (الف) زوم ۸ برابر تصویر اصلی، سطر (ب) روش درون‌یابی spline و سطر (ج) تصاویر الگوریتم سوپررزولوشن

گلسنر<sup>۱</sup> و همکاران در سال ۲۰۰۹ نیز یک یادگیری بدون نیاز به تصاویر HR پیشنهاد دادند. آن‌ها برای آموزش، از داده‌های تصویر ورودی در مقیاس‌های متفاوت استفاده نمودند. این رویکرد بر اساس فرض رخداد الگوهای خود متشابه در مقیاس‌های مختلف صورت پذیرفته است. در این روش یک مقیاس کوچک‌تر از تصویر اولیه تولید شده، سپس به ازای پنجره‌هایی با ابعاد معلوم (مثلاً  $7 \times 7$ ) از تصویر اولیه، موقعیت‌های شبیه در تصویر با مقیاس پایین‌تر جست‌وجو می‌شود. در گام بعدی، موقعیت‌های متناظر پنجره‌های شناسایی شده، در تصویر با مقیاس بزرگ‌تر به‌عنوان داده متناظر در روند آموزش یک شبکه افزایش وضوح مکانی انتخاب می‌شوند [۵۹].

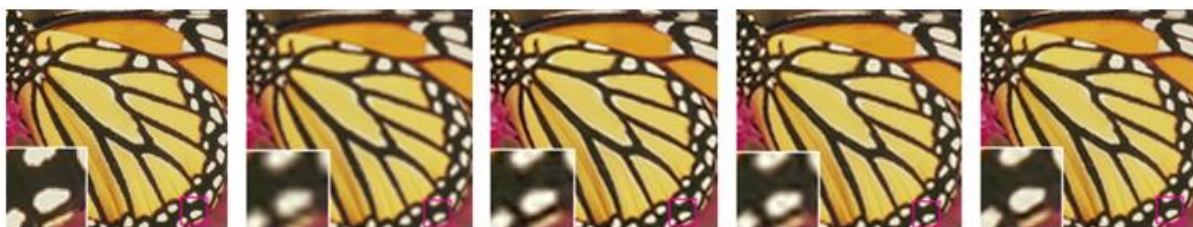
<sup>۱</sup> Daniel Glasner

در سال ۲۰۱۵ دانگ<sup>۱</sup> و همکاران نیز یک روش یادگیری عمیق مبتنی بر CNN<sup>۲</sup> برای بهبود وضوح مکانی به کمک تنها یک تصویر، پیشنهاد کردند. آن‌ها یک نگاهت انتها به انتها بین تصویر اولیه و تصویر LR تولیدشده از آن، در قالب یک شبکه CNN طراحی نمودند. در اولین لایه کانولوشنی این طراحی، مجموعه‌ای از نقشه‌های ویژگی از تصویر ورودی تولیدشده که در لایه دوم، این نقشه‌های ویژگی به صورت غیرخطی به نمایش‌هایی با وضوح بالا تبدیل می‌شوند. آخرین لایه مسئول ادغام نقشه‌های لایه‌های قبلی با هدف تولید یک تصویر با وضوح بالاتر است (شکل ۲-۲۰)). ایشان مدعی شدند که معماری CNN آموزش‌دیده با لایه‌های کانولوشنی متعدد و توابع فعال‌سازی غیرخطی و با بهره‌برداری از ویژگی‌های سطح پایین و بالا، نقشی کلیدی در فرایند بهبود وضوح مکانی دارند. نتایج مقایسه بصری برخی از نتایج شبکه ایشان با شبکه‌های رقیب در شکل (۲-۲۱) قابل مشاهده است [۶۰].



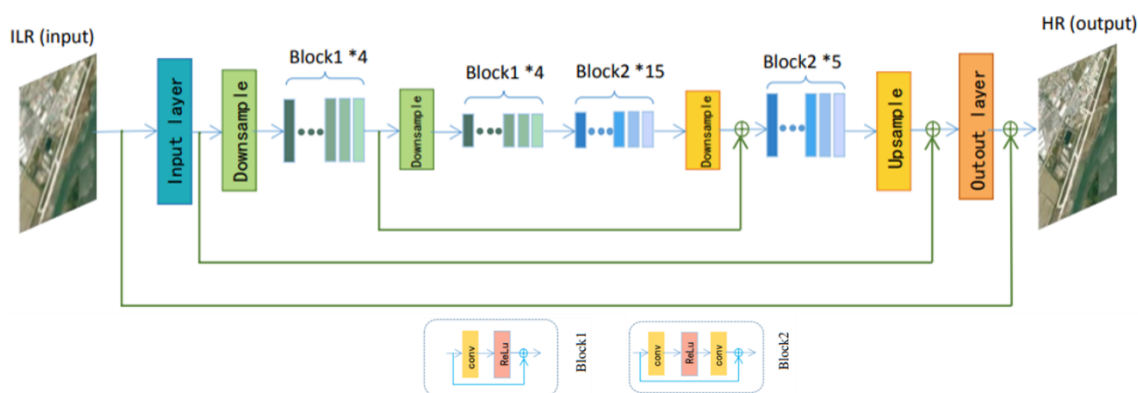
شکل ۲-۲۰: ساختار شبکه عمیق کانولوشنی استفاده شده در این مقاله

Chao Dong<sup>۱</sup>  
Convolutional Neural Networks<sup>۲</sup>



شکل ۲-۲۱: تصاویر از چپ به راست به ترتیب تصویر اورجینال، نتیجه حاصل از الگوریتم bicubic، نتایج حاصل از شبکه‌های SC، KK و نتیجه شبکه معرفی شده SRCNN می‌باشند.

در سال 2018 وین<sup>۱</sup> و همکاران برای طبقه‌بندی تصاویر سنجش از دوری، به علت وجود محدودیت‌های وضوح مکانی از تکنیک‌های سوپررزولوشن برای افزایش توان تفکیک مکانی استفاده کردند. در این راستا از یک الگوریتم جدید به نام شبکه عمیق متصل به حافظه (DMCN<sup>۲</sup>) استفاده نمودند. این شبکه با الهام گرفتن از مکانیسم مغز انسان، ارتباط حافظه محلی و جهانی را برای ترکیب جزئیات تصویر استفاده می‌کنند. برای کاهش زمان پردازش و کمتر شدن پارامترهای اضافی از نمونه‌برداری کاهشی استفاده نمودند. ساختار شبکه پیشنهادی ایشان در شکل (۲-۲۲) قابل مشاهده می‌باشد. نتایج مقایسه بصری و عددی این روش گویای برتری در مقایسه با روش ای رقیب می‌باشد (شکل (۲-۲۳)) [۶۱].



شکل ۲-۲۲: معماری شبکه عمیق پیشنهادی DMCN

Wen<sup>۱</sup>

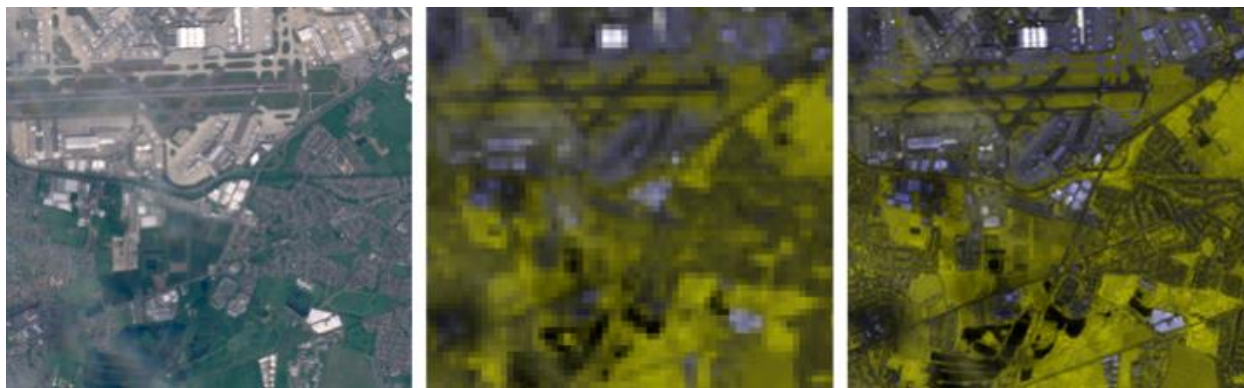
Deep Memory Connected Network<sup>۲</sup>



شکل ۲-۲۳: تصاویر از چپ به راست به ترتیب از تصویر دوم، تصویر وضوح بالا، نتیجه روش bicubic، نتیجه شبکه SRCNN، نتیجه شبکه vdsr، نتیجه شبکه DMCN

در سال ۲۰۱۸ چاریس لاناراس<sup>۱</sup> و همکاران از تکنیک‌های سوپرزولوشن تصویر برای تبدیل باندهای ۶۰ متری به ۲۰ متری (شکل (۲-۲۴)) و باندهای ۲۰ متری به ۱۰ متری (شکل (۲-۲۵)) تصاویر سنتینل-۲ استفاده نمودند. به این منظور آن‌ها از یک شبکه کانولوشنال انتها به انتها پیشرفته (CNN) برای افزایش وضوح تصاویر استفاده کردند. برای گسترش شبکه آموزش دیده، آن‌ها از داده‌های نمونه‌گیری شده در سطح گسترده‌ای از مکان‌های جغرافیایی استفاده نمودند تا شبکه‌ای را به دست آورند که در مناطق مختلف آب و هوایی و انواع پوشش زمین تعمیم می‌یابد و می‌تواند تصاویر دلخواه سنتینل-۲ را بدون نیاز به آموزش مجدد ارتقای مکانی بخشد. در ارزیابی‌های کمی (در مقیاس پایین‌تر، جایی که حقیقت زمینی در دسترس است)، این شبکه که آن را ۲DSen می‌نامند، تقریباً ۵۰٪ در معیار RMSE از بهترین رویکرد رقابتی بهتر عمل می‌کند. این روش در حالی که ویژگی‌های طیفی را بهتر حفظ می‌کند، نتایج بصری قانع‌کننده‌ای را در ۱۰ متر ارائه می‌دهد [۶۲].

<sup>۱</sup> Charis Lanaras



شکل ۲۴-۲: نتایج مربوط به تولید تصاویر ۶۰ متری از تصاویر ۲۰ متری. تصاویر از چپ به راست به ترتیب مربوط به تصویر ۲۰ متری واقعی، ترکیب سه باند ۶۰ متری و ترکیب سه باند ۲۰ متری تولیدشده توسط شبکه.



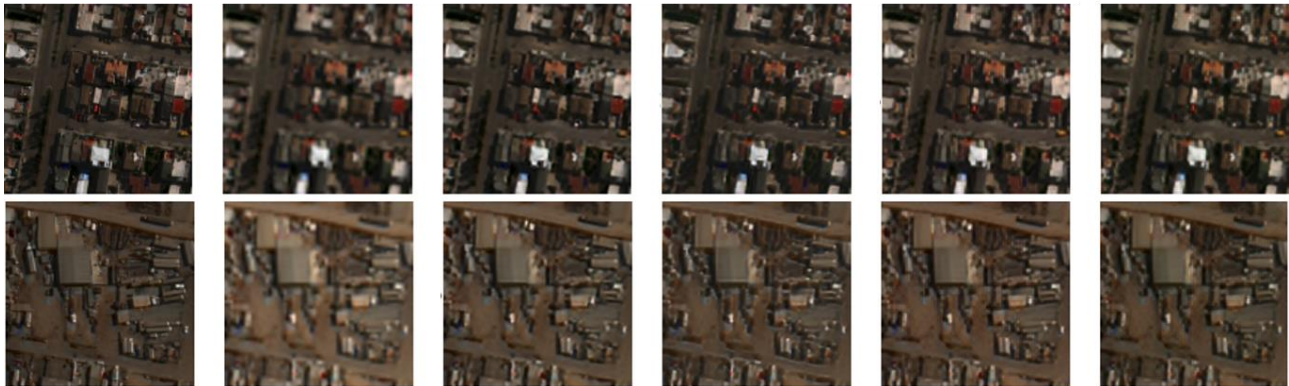
شکل ۲۵-۲: نتایج مربوط به تولید تصاویر ۱۰ متری از تصاویر ۲۰ متری. تصاویر از چپ به راست به ترتیب مربوط به تصویر ۱۰ متری واقعی، ترکیب سه باند ۲۰ متری و ترکیب سه باند ۱۰ متری تولیدشده توسط شبکه.

در سال ۲۰۱۹ هوک<sup>۱</sup> و همکاران، تکنیک‌های سوپرزولوشن تصویر مبتنی بر یادگیری عمیق با استفاده از شبکه عصبی کانولوشن (CNN) و شبکه متخاصم مولد (GAN<sup>۲</sup>) را برای افزایش چهار برابری وضوح تصاویر سنجش از دور

---

Hoque<sup>۱</sup>  
Generative Adversarial Network<sup>۲</sup>

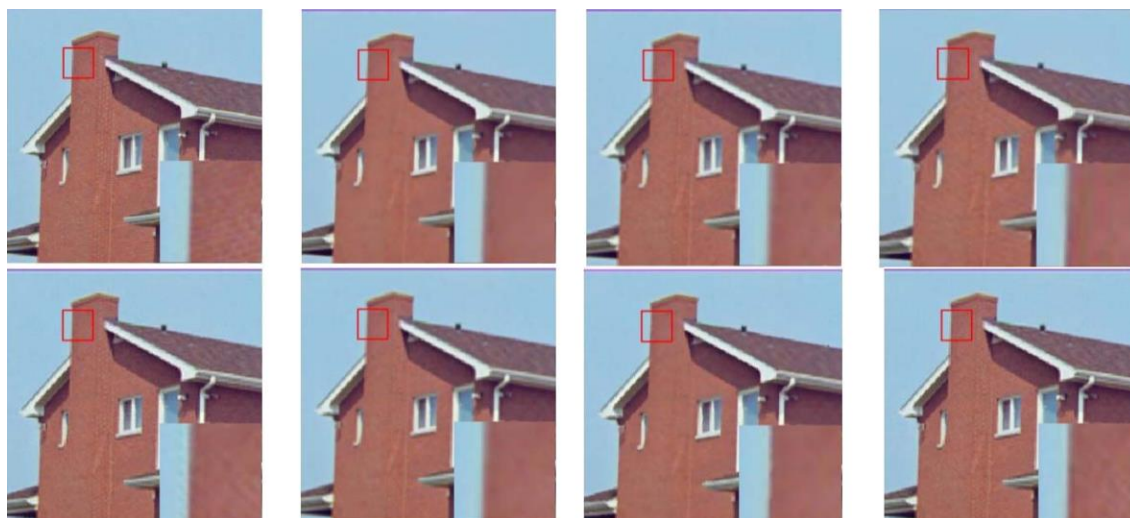
پیشنهاد دادند. نتایج تجربی نشان می‌دهد که مدل GAN عملکرد بهتری نسبت به مدل‌های CNN داشته است (شکل ۲-۲۶) [۶۳].



شکل ۲-۲۶: تصاویر از چپ به راست به ترتیب: تصویر HR، نتایج Bicubic، نتایج CNN-1، نتایج CNN-2، نتایج GAN-1، نتایج FGAN-2

در سال ۲۰۲۰ حبیبی و همکاران یک روش خود یادگیرنده برای افزایش وضوح تصاویر ارائه نمودند. در این روش دو تصویر LR و HR ساخته می‌شود. رابطه بین تصویر LR و HR توسط رگرسیون بردار پشتیبان (SVR<sup>1</sup>) یادگیری می‌شود. برای ایجاد تخمین بهتری از روشنایی تصویر HR، دو تصویر LR و HR آموزشی را بر اساس رنگ ناحیه، ناحیه بندی نموده و مدل‌های رگرسیونی را برای هر ناحیه به‌طور مجزا آموزش داده‌اند. از طرفی برای کاهش اثر تاری در لبه‌های تصویر ارتقا یافته، مدل‌های جداگانه‌ای برای یادگیری روشنایی لبه‌ها ارائه شده است. ویژگی‌های بکار رفته در یادگیری SVR، ضرایب تنک وصله‌ها در بازنمایی تنک و گرادیان وصله‌ها می‌باشد که برای هر کدام مدل رگرسیون جداگانه‌ای آموزش داده شده و خطای این مدل‌ها به کمک SVR مدل‌سازی می‌شود. نتایج نشان می‌دهد که روش پیشنهادی نسبت به سایر روش‌ها با توجه به معیارهای PSNR<sup>2</sup> و SSIM<sup>3</sup> عملکرد بهتری در ارتقا توان تفکیک مکانی دارد (شکل ۲-۲۷) [۶۴].

<sup>1</sup> Support Vector Regression  
<sup>2</sup> Peak Signal-to-Noise Ratio  
<sup>3</sup> Structural Similarity Index Measure



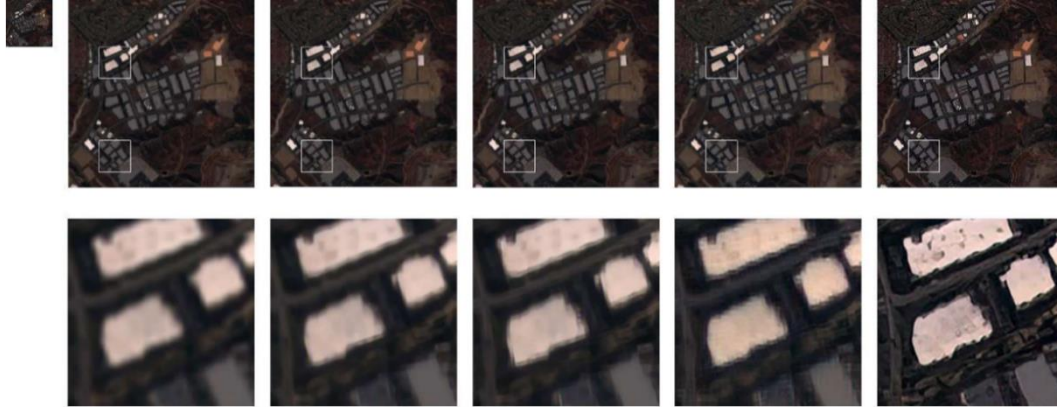
شکل ۲-۲۷: نتایج ارتقا ۲ برابری وضوح. به ترتیب از چپ به راست و بالا به پایین، تصویر هدف، روش SISR، روش Bicubic، روش LLE، روش ScSR، روش LsSR، روش SINE و روش پیشنهادی.

در سال ۲۰۲۱ لی<sup>۱</sup> و همکاران به کمک آموزش یک شبکه عمیق به افزایش وضوح چهار برابری باندهای ۱۰ متری سنجنده سنتینل-۲ پرداختند. مدل استفاده شده توسط آن‌ها DKN-SR-GAN نام دارد که از تخمین هسته تخریب و تزریق نویز برای ساخت مجموعه داده‌ای از جفت تصویر با وضوح پایین و وضوح بالا استفاده می‌کند. آن‌ها از تصاویر نمونه برداری کاهشی شده سنجنده وردیو-۲ به عنوان تصویر مرجع ارتقا وضوح یافته استفاده نموده‌اند. نتایج به دست آمده (شکل (۲-۲۸)) بیانگر عملکرد بهتر شبکه نسبت به روش‌های دیگر در معیارهای ارزیابی کیفیت می‌باشد [۶۵].

---

<sup>۱</sup> Li





شکل ۲-۲۸: در سمت چپ بالا تصویر سنتینل-۲ را مشاهده می‌شود و تصاویر بزرگ‌تر به ترتیب از چپ به راست، تصاویر سوپرزولوشن شده به روش bicubic، شبکه EDSR8، شبکه RCAN، شبکه SRGAN، شبکه DKN-SR-GAN می‌باشند.

## ۲-۴- جمع بندی

در این فصل به بررسی و طبقه‌بندی تحقیقات انجام‌شده در زمینه بهبود وضوح تصاویر و تحلیل نقاط قوت و ضعف آن‌ها پرداخته شد. الگوریتم‌های مطرح شده در این زمینه در دو گروه روش‌های مبتنی تلفیق (بخش ۲-۲)، روش‌های سوپرزولوشن (بخش ۲-۳)، مطرح شدند که در توصیف هر گروه ابتدا مفاهیم پایه مربوطه معرفی شد و سپس تحقیقات پیشین در آن زمینه مورد بررسی قرار گرفت.

## فصل ۳: توسعه نظری راهکار پیشنهادی برای حل مساله

### ۳-۱- مقدمه

در این فصل به معرفی روشی دو مرحله ای جهت بهبود توان تفکیک مکانی تصاویر سنجش از دوری در عین حفظ توان تفکیک طیفی در زمان فقدان تصویری با توان تفکیک مکانی بالاتر (پانکروماتیک) پرداخته شده است. در ادامه به توضیحات کاملی از جزئیات این فرآیند پرداخته شده است.

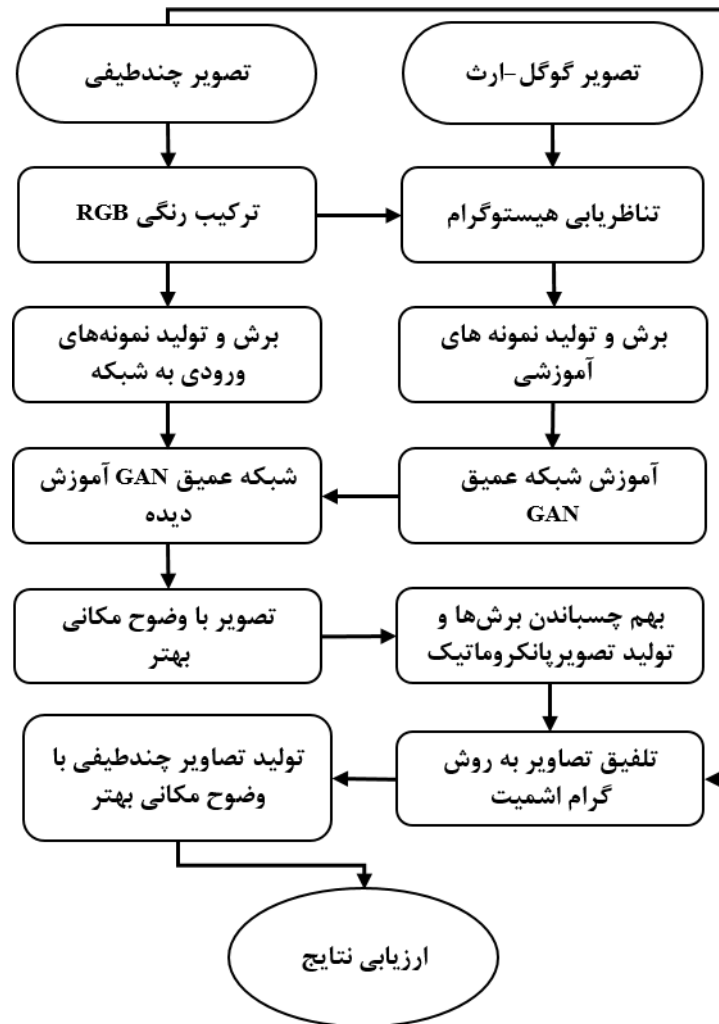
### ۳-۲- روش پیشنهادی

امروزه شبکه‌های عصبی آموزش یافته به کمک تکنیک‌های یادگیری عمیق در حوزه‌های مختلفی از پردازش تصاویر سنجش از دوری استفاده می‌شوند. مطابق با منابع مرور شده در فصل (۲)، روش‌های متنوعی به منظور ارتقاء وضوح مکانی تصاویر سنجش از دوری به کمک تکنیک‌های یادگیری عمیق توسعه یافته است. در تمامی رویکردهای مبتنی بر یادگیری عمیق دسترسی به یک پایگاه داده غنی از نمونه‌های آموزشی پیش شرط موفقیت این روش‌ها بیان شده است [۶۶]. از منظر کاربردی، دستیابی به چنین محتوای اطلاعاتی با چالش‌های اجرایی روبروست. به عبارتی آموزش شبکه ای که توانایی بهبود کیفیت مکانی تصاویر را داشته باشد نیازمند در اختیار داشتن تعداد بسیار زیادی تصویر HR و LR می‌باشد که این مسئله یک چالش محسوب می‌شود. بعنوان مثال در سنجنده ماهواره‌ای سنتینل-۲ که فاقد باند پانکروماتیک است؛ تامین داده‌های آموزشی با محتوای طیفی مشابه و توان تفکیک مکانی بالاتر توأم با مشکلات و هزینه‌های اجرایی است. بکارگیری تصاویر ماهواره‌ای HR بعنوان تصاویر متناظر، عدم تطابق در باندهای طیفی و چالش‌های هم‌مرجع‌سازی را بدنبال دارند. به عبارت بهتر، عدم مشابهت موقعیت، عرض و تابع

پاسخ طیفی سنجنده‌های چندطیفی امکان تامین داده‌های آموزشی چندطیفی در دو نسخه HR و LR را با محدودیت مواجه می‌سازد. از این رو در این پژوهش ایده‌ای کم هزینه برای تولید تصاویر HR چندطیفی پیشنهاد شده است.

گوگل-ارث یک پایگاه داده غنی از تصاویر ماهواره‌ای بوده که سراسر پهنه خشکی‌های زمین را در بر دارد. این منبع اطلاعاتی مبنای اجرای کاربردهای مختلفی بوده که متخصصین نقشه‌برداری به ندرت از آن بی‌بهره بوده‌اند. تاکنون راهکارهای متنوعی به منظور ارتقاء بهره‌وری از تصاویر این پایگاه داده توسعه یافته است. ایجاد نرم‌افزارهای دانلود و زمین‌مرجع‌سازی دقیق تصاویر گوگل مانند stitchMapper، Google-Earth downloader و SasPlanet نمونه نرم‌افزارهایی تلقی شده که با هدف دانلود و بهبود دقت مکانی تصاویر گوگل بر پایه روش‌های پردازش تصویری توسعه یافته‌اند. به کمک این نرم‌افزارها و پایگاه داده گوگل-ارث می‌توان تصاویر ماهواره با وضوح حدوداً ۱ متر تهیه نمود و با نمونه برداری کاهشی به وضوح دلخواه دست یافت. ایده این تحقیق، آموزش عمیق شبکه‌های عصبی بر پایه بکارگیری تصاویر گوگل-ارث است. در این روش، تصاویر گوگل-ارث مربوط به منطقه مورد مطالعه پس از تنظیم توان تفکیک مکانی آن متناسب با تصاویر ورودی و خروجی مورد انتظار، برای آموزش شبکه‌های عمیق استفاده می‌شود. این تصاویر صرفاً در محدوده مکانی تصاویر چندطیفی و با کمترین اختلاف زمانی ممکن انتخاب می‌شوند. پس از آموزش یک شبکه عصبی به روش یادگیری عمیق بر پایه تصاویر گوگل-ارث (پس از اعمال اصلاحات هندسی و رادیومتریکی)، از این شبکه‌ها برای بهبود وضوح تصاویر چندطیفی استفاده می‌شود. محصول این مرحله یک تصویر با وضوح مکانی بالا بوده و از وضوح طیفی مناسبی برخوردار نیست. به همین منظور از این تصویر به عنوان تصویر پانکروماتیک در ادامه استفاده خواهد شد. در گام دوم، تصاویر با وضوح مکانی بالاتر بدست آمده از نتایج یادگیری عمیق (تصویر پانکروماتیک تولید شده) با باندهای طیفی تصاویر چندطیفی از طریق روش‌های مرسوم (فصل ۲) تلفیق می‌شوند. محصول نهایی در این روند بدون نیاز به تصاویر با وضوح مکانی و

طیفی بالاتر و صرفاً با بهره‌گیری از تصاویر منتشر شده در بستر گوگل-ارث تولید می‌شوند. روندنمای ارائه شده در شکل (۱-۳) روند کلی روش پیشنهادی را نشان می‌دهد.



شکل ۳-۱ روند نمای کلی روش پیشنهادی

مطابق با روند نمای شکل (۱-۳)، به ازای هر تصویر چندطیفی، تصاویر HR متناظر از بستر گوگل-ارث استخراج شده و در دو نسخه متناسب با توان تفکیک مکانی تصویر چندطیفی و تصویر ارتقاء یافته مکانی مورد انتظار، نمونه‌برداری کاهشی می‌گردد. قبل از این اقدام شباهت رادیومتریکی تصاویر گوگل به تصاویر چندطیفی از طریق

روش تناظریابی هیستوگرام افزایش می‌یابد. این پروسه با باند های RGB تصویر چند طیفی صورت می‌پذیرد. پس از تنظیم وضوح مکانی و تناظریابی رادیومتریکی تصاویر مستخرج از سامانه گوگل-ارث، نمونه‌های آموزشی و ارزیابی به منظور آموزش عمیق یک شبکه عصبی تخصصی تولید می‌شود. این شبکه قادر خواهد بود که توان تفکیک مکانی را تا ۴ برابر ارتقاء بخشد. بازیابی محتوای طیفی در این روند به دلیل فقدان مابه‌ازای طیفی مناسب در روند آموزش میسر نبوده و این روند صرفاً با هدف تولید محتوای مشابه با تصاویر پانکروماتیک با وضوح مکانی بالا به اجرا می‌رسد. از این رو، شبکه آموزش یافته ترکیب رنگی واقعی از تصاویر چندطیفی را به عنوان ورودی دریافت کرده و سپس تصاویر با وضوح مکانی بالاتر را تولید می‌سازد که این تصاویر به لحاظ طیفی اطلاعات صحیحی ندارند. در گام بعد با میانگین‌گیری از باند های تصویر تولید شده باند پانکروماتیکی تولید می‌شود که به لحاظ هندسی بسیار شبیه به تصویر چند طیفی اولیه می‌باشد. این تصویر پس از تبدیل به نسخه تک‌باندی، با تصاویر چندطیفی اولیه به کمک یکی از روش های مرسوم (فصل ۲) تلفیق شده و محصول ارتقاء یافته را که تصویری با وضوح مکانی بالاتر در عین حفظ اطلاعات طیفی است، تولید می‌کنند. در ادامه این فصل هر یک از اجزای این روند در قالب زیربخش‌های مجزا به تفصیل تشریح شده‌اند.

### ۳-۳- شبکه عمیق GAN

یکی از شبکه های عصبی مهم و معروف در یادگیری ماشین، شبکه عصبی <sup>۱</sup>GAN است که این شبکه اولین بار در سال ۲۰۱۴ توسط گودفلو<sup>۲</sup> معرفی شد [۶۷]. در ادامه توضیح مختصری از شبکه عمیق GAN ارائه شده و سپس شبکه عمیق پیشنهادی تشریح شده است.

---

<sup>۱</sup> Generative Adversarial Network

<sup>۲</sup> Goodfellow

## ۳-۳-۱- معرفی شبکه عمیق GAN

شبکه های عصبی GAN مدل های مولدی<sup>۱</sup> هستند که داده هایی جدید شبیه داده های آموزشی تولید می کنند. شبکه های عصبی GAN می توانند تصاویری مانند چهره انسان تولید کنند که کاملاً ساختگی هستند. این شبکه ها از دو مدل با نام های مولد<sup>۲</sup> و متمایزگر<sup>۳</sup> تشکیل شده اند که این مدل ها به صورت خودکار شاخص های موجود در داده های ورودی را استخراج کرده و مورد آموزش قرار می گیرند. به طور کلی مولد و متمایزگر، برای بررسی، ضبط و تکرار تغییرات درون مجموعه داده با یکدیگر رقابت می کنند. می توان از GAN ها برای تولید نمونه های جدیدی که به طرز قابل قبولی از مجموعه داده اصلی قابل تهیه هستند، استفاده کرد. [۶۷]

مولد: یک شبکه عصبی است که وظیفه آن تولید داده های جعلی است تا متمایزگر بتواند به کمک آنها آموزش ببیند. مولد آموزش می بیند تا داده هایی قابل قبول و بسیار شبیه به واقعیت تولید کند. داده های تولید شده توسط مولد برای متمایزگر، نمونه های منفی<sup>۴</sup> آموزشی به حساب می آیند. به طور کلی مولد، یک بردار نویز تصادفی با طول ثابت را به عنوان ورودی دریافت کرده و سپس یک نمونه تولید می کند. هدف اصلی مولد فریب دادن متمایزگر است. به عبارت بهتر هدف مولد تولید داده ای است که متمایزگر آن را به عنوان داده واقعی یا مثبت<sup>۵</sup> دسته بندی کند. بخشی از شبکه GAN که وظیفه آموزش مولد را به عهده دارد شامل موارد زیر است:

۱- بردار نویز ورودی

۲- شبکه مولد، که وظیفه تولید یک داده از نویز ورودی را به عهده دارد.

۳- شبکه متمایزگر، که داده های تولید شده توسط مولد را به دو دسته مثبت و منفی تقسیم می کند.

---

<sup>۱</sup> Generative Models

<sup>۲</sup> Generator

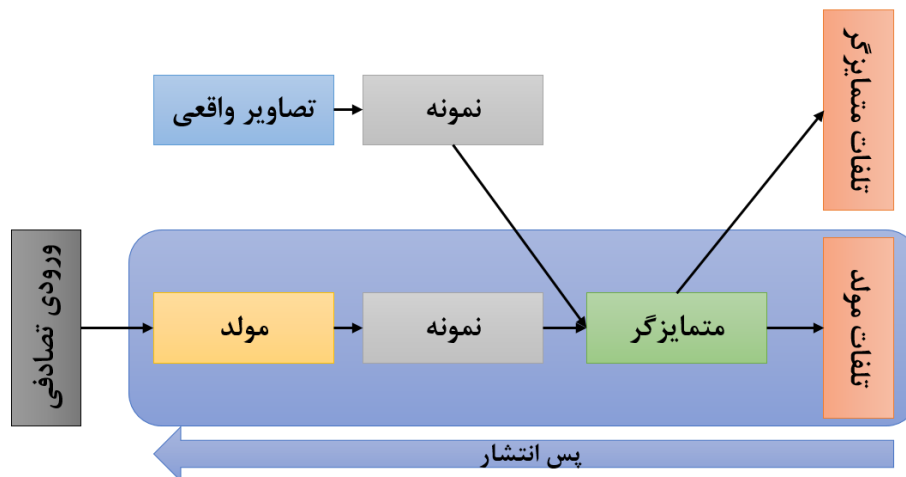
<sup>۳</sup> Discriminator

<sup>۴</sup> negative

<sup>۵</sup> positive

۴- تلفات مولد که بر اساس میزان اشتباه شبکه مولد در تولید داده محاسبه میشود که کمینه شدن این مقدار بیانگر واقعی بودن داده تولید شده توسط مولد میباشد.

از الگوریتم پس انتشار<sup>۱</sup> نیز در راستای تنظیم هر وزن در جهت مناسب با محاسبه تاثیر وزن بر خروجی استفاده می شود. همچنین از این روش برای به دست آوردن گرادینان استفاده می شود و این گرادینان ها می توانند به تغییر وزن های مولد کمک کرده و آنها را بهبود دهند (شکل (۲-۳)).

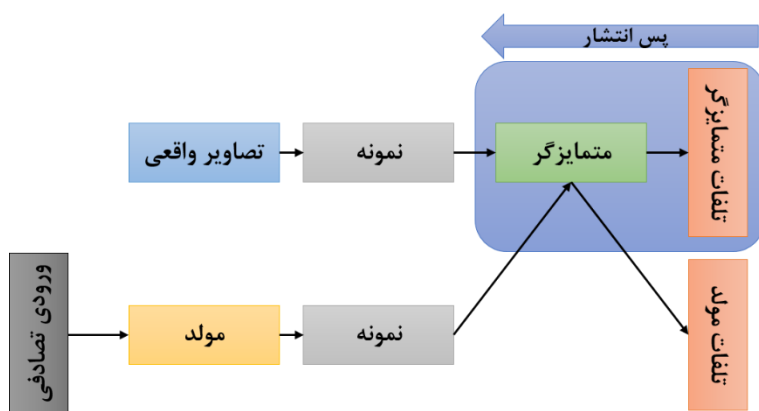


شکل ۲-۳: روند نمای آموزش مولد

تمتایزگر: وظیفه متمایزگر تشخیص داده های واقعی از داده های جعلی است؛ به عبارت بهتر بر روی داده های تولید شده توسط مولد با در دست داشتن داده های واقعی، برچسب گذاری میکند. آموزش متمایزگر توسط دو منبع داده آموزشی میباشد: ۱) نمونه های واقعی داده ها مانند تصاویر واقعی طبیعت، انسان، اسکناس های ارزی و غیره که توسط متمایزگر به عنوان نمونه های مثبت یا واقعی در طول آموزش استفاده میشوند، ۲) نمونه های غیر واقعی تولید شده توسط مولد که به عنوان نمونه های منفی یا جعلی در طول آموزش مورد استفاده قرار میگیرند.

<sup>۱</sup> backpropagation

در طول روند آموزش، متمایزگر به دو تابع اتلاف<sup>۱</sup> منتهی میشود که یکی برای آموزش مولد و دیگری برای آموزش متمایزگر قرار داده شده است. در پروسه آموزش شبکه متمایزگر، از تلفات مولد چشم پوشی شده و فقط از تلفات متمایزگر استفاده می‌شود. به این صورت که متمایزگر در حین فرآیند آموزش، داده‌های واقعی و داده‌های جعلی دریافتی از مولد را دسته بندی می‌کند. سپس در صورت دسته بندی نادرست یک نمونه داده واقعی به عنوان نمونه جعلی یا برعکس، مقدار تابع اتلاف زیاد میشود. متمایزگر، وزن‌های خود را با پس انتشار از تلفات متمایزگر در طول شبکه خود، به روز می‌کند (شکل (۳-۳)).



شکل ۳-۳: روند نمای آموزش متمایزگر

### ۳-۳-۲- معماری های استفاده شده در شبکه GAN پیشنهادی

در شبکه GAN پیشنهادی از دو معماری معروف جهت بهبود عملکرد شبکه استفاده شده است که در ادامه به معرفی این دو معماری به طور مختصر پرداخته شده است.

<sup>۱</sup> loss function



معماری رزنت<sup>۱</sup> یا شبکه باقیمانده<sup>۲</sup> یکی از معماری های معروف در شبکه های عمیق است. این شبکه اولین بار توسط شیوکینگ رن<sup>۳</sup> و همکارانش در سال ۲۰۱۵ معرفی شد که توانست برنده چالش ILSVRC در سال ۲۰۱۵ شود. این معماری به علت طراحی خاصش به شبکه عمیق اجازه میدهد که بتواند تا عمق ۱۵۰ لایه را پردازش کند در حالی که شبکه های عمیق پیش از آن به دلیل مشکل محوشدگی گرادیان<sup>۴</sup> دچار مشکل میشدند و توانایی پردازش لایه های بسیار زیاد را نداشتند. راه حل این معماری جهت عمیق تر کردن شبکه استفاده از اتصالات میانبر<sup>۵</sup> یا اتصالات اضافی<sup>۶</sup> میباشد. با توجه به شکل (۳-۴) اتصالاتی در این شبکه موجود است که از یک یا چند لایه عبور میکند و لایه های میانی را نادیده میگیرد. به عبارت بهتر یک اتصال دو لایه دور از هم را با میانبر زدن و به صورت مستقیم بهم وصل میکند [۶۸].

---

<sup>۱</sup> ResNet

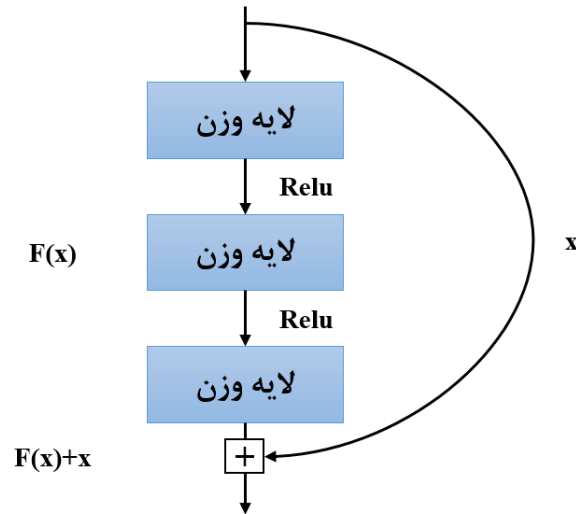
<sup>۲</sup> Residual Network

<sup>۳</sup> Shaoqing Ren

<sup>۴</sup> Vanishing Gradient

<sup>۵</sup> Skip Connections

<sup>۶</sup> Residual Connections

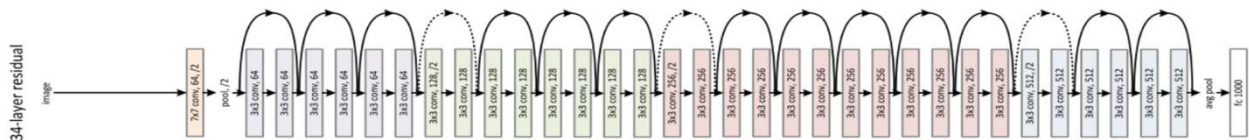


شکل ۳-۴: نحوه اتصال میانبر در معماری ResNet

با توجه به شکل (۳-۴)، اولین تفاوتی که به نظر میرسد آن است که یک اتصال مستقیم وجود دارد و این اتصال از لایه‌هایی میانی عبور کرده است. به این اتصال، اتصال میانبر گفته میشود و عنصر اصلی بلوک‌های باقیمانده است. به علت وجود چنین اتصالی، نتیجه لایه مشابه نتایج لایه‌های شبکه‌های دیگر نیست. در این اتصال ورودی  $x$  در وزن لایه‌ها ضرب شده و همچنین مقدار بایاس به آن اضافه می‌شود. سپس این عبارت از تابع فعالسازی  $F(x)$  عبور می‌کند و در نهایت خروجی آن نتیجه میشود که با  $H(x)$  نمایش داده شده است که به صورت رابطه (۳-۱) تعریف میشود.

$$H(x) = F(x) + x \quad \text{رابطه (۳-۱)}$$

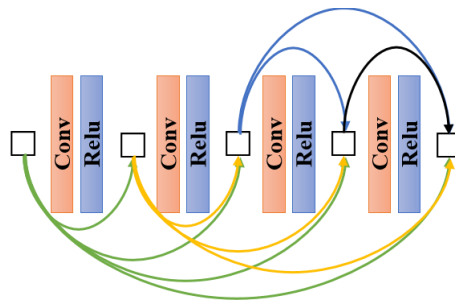
شکل (۳-۵) یک شبکه ۳۴ لایه‌ای از معماری ResNet را به نمایش گذاشته است [۶۸].



شکل ۳-۵: معماری ۳۴ لایه‌ای شبکه ResNet

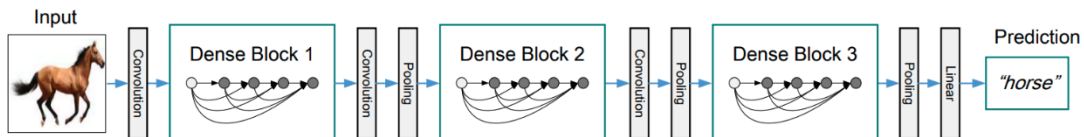
## ۳-۲-۲-۲- معماری DensNet

معماری DensNet یک معماری الگو گرفته از شبکه ResNet است. در این معماری به جای استفاده از اتصال میانبر از لایه قبل به لایه های بعد، از اتصالات تمام لایه ها بهم استفاده شده است. به طور کلی در شبکه های عصبی کانولوشن لایه های ابتدایی ویژگی های سطح پایین مثل لبه ها را استخراج می کنند و لایه هایی که در انتهای این زنجیره قرار گرفتند ویژگی های سطح بالا مانند بافت ها را استخراج می کنند. در برخی موارد ممکن است در عملیات طبقه بندی یک کلاس از ویژگی های سطح پایین سود بیشتری ببرد. در معماری DensNet با توجه به متصل بودن لایه هایی ابتدایی به لایه های انتهایی شبکه می تواند آموزش ببیند که برای کلاس مورد نظر فقط از ویژگی های سطح پایین یا از ترکیب ویژگی های سطح پایین و بالا استفاده کند. همچنین هر لایه دسترسی مستقیم به مشتق تابع هزینه و ورودی اصلی شبکه دارد این مورد به آموزش معماری های عمیق تر کمک میکند. شکل (۳-۶) بیانگر نحوه اتصال لایه ها به یکدیگر در معماری DensNet را نشان میدهد [۶۹].



شکل ۳-۶: نحوه اتصالات در بلوک های DensNet

نمونه ای از نحوه قرار گیری بلاک های DensNet در یک شبکه عمیق در شکل (۳-۷) نشان داده شده است [۷۰].



شکل ۳-۷: آرایش قرار گیری بلاک های DensNet در یک شبکه عمیق کلاسه بندی

### ۳-۳-۳ - شبکه عمیق GAN پیشنهادی

مطابق با محتوای مندرج در روندنمای روش پیشنهادی شکل (۳-۱) تولید تصاویر با توان تفکیک مکانی بالا بمنظور تلفیق با تصاویر چندطیفی به اجرا رسیده است. برای این منظور از یک شبکه عمیق (Real-ESRGAN) استفاده شده است که یک الگوریتم سوپرزولوشن مبتنی بر یادگیری عمیق است و می تواند وضوح مکانی تصاویر با وضوح مکانی پایین را افزایش دهد [۷۱]. Real-ESRGAN یک نسخه پیشرفته از ESRGAN [۷۲] است که این شبکه یک الگوریتم قدرتمند سوپرزولوشن بر اساس معماری GAN [۷] است. ساختار کلی Real-ESRGAN معماری GAN بوده که در داخل آن از اتصالات ResNet و DensNet استفاده شده است. Real-ESRGAN برای افزایش وضوح مکانی چهار برابری تصویر از یک فرآیند دو مرحله ای استفاده می کند. به طوریکه ابتدا با استفاده از شبکه مولد یک تصویر اولیه با وضوح مکانی بالا تولید می کند. سپس از شبکه متمایزگر برای مقایسه تصویر بدست آمده با تصویر مرجع دارای وضوح مکانی بالا و شناسایی تفاوت ها استفاده می کند. سپس از این تفاوت ها برای به روز رسانی شبکه مولد استفاده می شود تا تفاوت بین تصویر تولید شده و تصویر مرجع را به حداقل برساند.

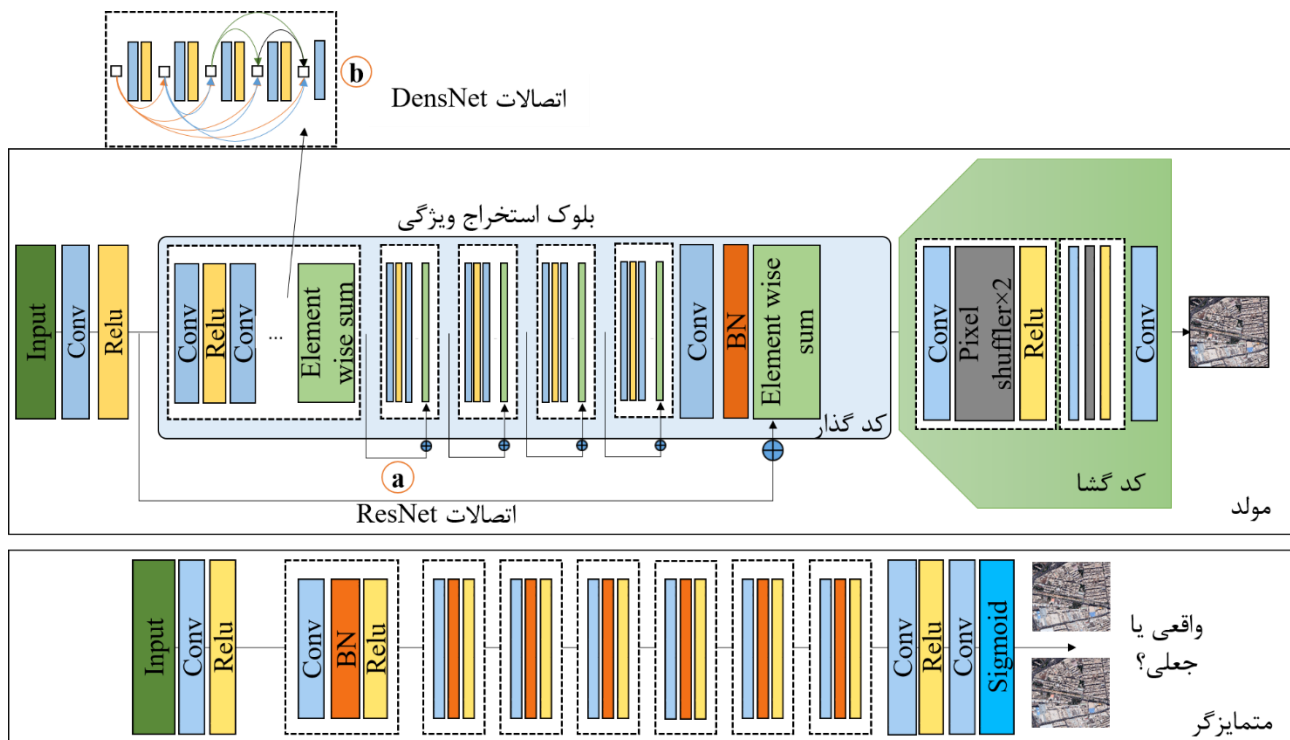
شبکه مولد پیشنهاد شده دارای دو بخش کدگذار<sup>۱</sup> و کدگشا<sup>۲</sup> است. بخش کدگذار شبکه مولد از تکنیکی به نام یادگیری باقیمانده استفاده می کند که به شبکه اجازه می دهد به جای تلاش مستقیم برای تولید تصویر با وضوح مکانی بالا، باقیمانده و یا تفاوت بین تصاویر با وضوح مکانی پایین و بالا را آموزش ببیند. این رویکرد می تواند عملکرد کلی شبکه را بهبود بخشد و به کاهش میزان داده های آموزشی مورد نیاز کمک کند [۷۱]. شبکه مولد از بلوک های شبکه DensNet [۶۹] تشکیل شده است که این بلوک ها مطابق با معماری باقیمانده به کار رفته در شبکه ResNet [۷۳] به هم متصل شده اند. به این صورت که یک اتصال مستقیم بین ورودی بلوک هر DensNet و خروجی آن در هر بار اعمال این بلوک ایجاد شده است (a در شکل (۳-۸)). هر بلوک DensNet خود از اعمال توالی

---

<sup>۱</sup> Encoder

<sup>۲</sup> Decoder

لایه‌های Conv، ReLU و Conv در دفعات متفاوت تشکیل شده است که در آن خروجی یک بار اعمال این توالی به نتایج حاصل از اعمال تمامی دیگر توالی‌ها متصل است (b در شکل (۳-۸)). در بخش کدگذاری شبکه مولد با اعمال لایه‌های Conv، pixel-shuffler و ReLU سعی می‌شود که با بهره‌گیری از ویژگی‌های بدست آمده از بخش کدگذاری شامل ویژگی‌های سطح پایین مانند لبه‌ها و بافت‌ها تا ویژگی‌های سطح بالا مانند عوارض، ویژگی‌هایی موثر با وضوح مکانی ۲ برابر تولید شود. در ادامه با تکرار فرایند مطرح شده، وضوح مکانی ویژگی‌ها بار دیگر دوبرابر شده و نهایتاً با اعمال یک لایه Conv تعداد کانال‌های ویژگی‌ها کاهش یافته و تصویر نهایی با وضوح مکانی ۴ برابر بدست می‌آید. شکل (۳-۸).



شکل ۳-۸: معماری شبکه عمیق Real-ESRGAN

در شبکه تفکیک کننده با اعمال لایه‌های Conv، BN و ReLU و بهره‌گیری از رابطه (۳-۲) سعی می‌شود که واقعی یا جعلی بودن تصویر تولید شده توسط شبکه مولد تعیین گردد.

$$\begin{cases} D(I_f, I_f) = \sigma(C(I_r) - E[C(I_f)]) \rightarrow 1 \\ D(I_f, I_f) = \sigma(C(I_f) - E[C(I_r)]) \rightarrow 0 \end{cases} \quad \text{رابطه (۲-۳)}$$

در رابطه (۲-۳)  $I_r$  تصویر واقعی،  $I_f$  تصویر جعلی،  $\sigma$  تابع سیگموئید،  $E$  تابع میانگین و  $C$  خروجی شبکه تفکیک کننده می باشد. اگر  $D$  به ۱ میل کند، نشان دهنده واقعی بودن تصویر تولید شده توسط مولد و اگر به ۰ میل کند نشان دهنده جعلی بودن تصویر تولید شده توسط مولد است. یکی از ویژگی های کلیدی Real-ESRGAN تابع اتلاف ادراکی آن است که برای بهینه سازی شبکه مولد استفاده می شود. این تابع اتلاف طراحی شده است تا نه تنها تفاوت بین تصاویر واقعی تولید شده و تصاویر حقیقی، بلکه شباهت ادراکی آنها را نیز اندازه گیری کند. این کار برای اطمینان حاصل کردن از یکسان بودن کیفیت بصری تصاویر تولید شده با تصاویر حقیقی است.

شبکه عمیق مذکور توسط مجموعه داده های متعددی آموزش دیده و نسخه آموزش یافته آن در آدرس (<https://github.com/xinntao/Real-ESRGAN>) در دسترس است. این شبکه عصبی در دو نسخه اولیه (بدون آموزش مجدد) و آموزش یافته تقویتی (به کمک داده های محدود از منطقه مورد نظر) در روند بازسازی تصاویر با وضوح بهتر استفاده شده است.

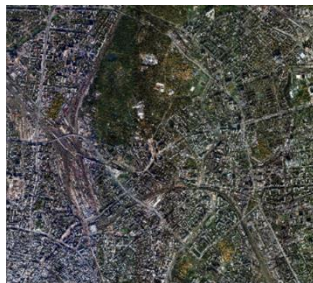
### ۳-۴ - آماده سازی نمونه های آموزشی برای یادگیری عمیق

به کارگیری نمونه های آموزشی مناسب پیش شرط موفقیت در عملکرد یک شبکه عصبی مصنوعی است. مطابق با آنچه بدان اشاره شد؛ در این تحقیق تصاویر گوگل-ارث برای آموزش شبکه عصبی Real-ESRGAN انتخاب شده اند. این موضوع باهدف کاهش هزینه های مربوط به تأمین داده های آموزشی صورت گرفته است. همچنین به علت گستردگی پوشش سامانه داده گوگل-ارث، محدودیتی از بابت تهیه داده آموزشی در نواحی مختلف جهان وجود ندارد. با این حال برای بهتر شدن نتایج شبکه آموزش دیده، دو پیش پردازش هندسی و رادیومتریکی به منظور انطباق بیشتر داده های آموزشی با تصاویر چند طیفی به اجرا رسیده است.

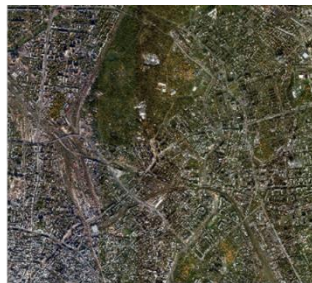
به دلیل ماهیت موزائیک شده تصاویر گوگل و تمایزات رادیومتریکی ناشی از صحنه و سنجنده، محتوای رادیومتریکی تصاویر گوگل دارای تنوع و ناهمگونی است. از این رو، پس از استخراج پهنه جغرافیایی مشابه با تصویر چند طیفی از بستر گوگل-ارث، محتوای رادیومتریکی آن‌ها با باندهای رنگی مرئی از تصویر چند طیفی شبیه شدند. همچنین مشابه شدن اطلاعات طیفی تصاویر آموزشی گوگل-ارث با تصاویر چند طیفی سنجنش از دوری باعث بهبود عملکرد این شبکه در افزایش وضوح مکانی تصاویر سنجنش از دوری خواهد شد. این اقدام از طریق اجرای روش تناظر یابی هیستوگرام صورت گرفته است. تناظر یابی هیستوگرام یک تبدیل نقطه‌ای بوده که شباهت توزیع درجات خاکستری تصویر منبع (گوگل) را به تصویر مرجع (چند طیفی) افزایش می‌دهد [۷۴]. رابطه (۳-۳) تابع تبدیل در فرایند تناظر یابی هیستوگرام است.

$$y = f_r^{-1}(f_s(x)) \quad \text{رابطه (۳-۳)}$$

در رابطه (۳-۳)،  $f_s$  تابع هیستوگرام تجمعی نرمال شده از تصویر منبع،  $f_r^{-1}$  تابع معکوس هیستوگرام تجمعی نرمال شده از تصویر مرجع،  $x$  درجه خاکستری تصویر منبع و  $y$  درجه خاکستری اصلاح شده پس از تناظر یابی هیستوگرام هستند شکل (۳-۹).



تصویر اولیه گوگل ارث

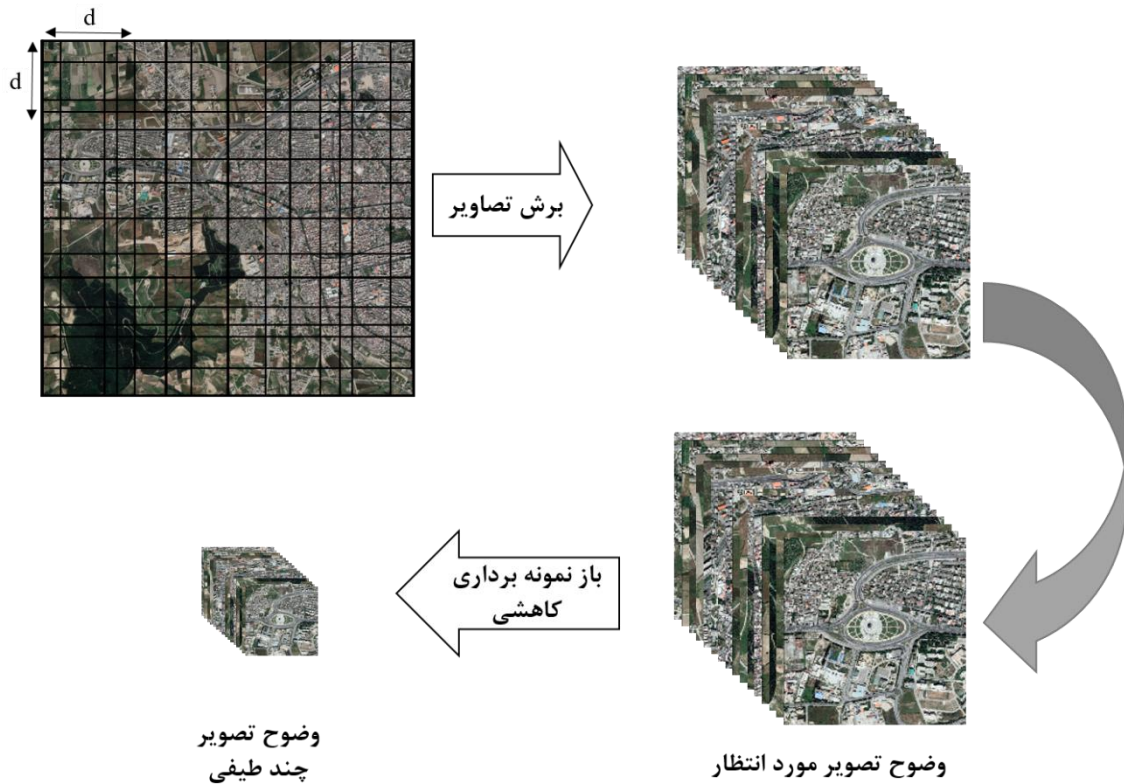


تصویر سه باند RGB سنتینل-۲ تصویر گوگل ارث حاصل از تناظر یابی هیستوگرام



شکل ۳-۹: سمت چپ به علت ماهیت موزائیک شده تصاویر گوگل-ارث با سنجنده های مختلف دچار تنوع و ناهمگونی رنگ شده است که در شکل وسط این چالش به کمک تناظر یابی هیستوگرام با تصویر سه بانده مرئی سنتینل-۲ کاهش یافته است. با توجه به عملکرد شبکه عصبی Real-ESRGAN در افزایش چهار برابری وضوح تصاویر و ضرورت تأمین محتوای تصویری مشابه با تصاویر چند طیفی در روند آموزش، وضوح تصاویر مستخرج از گوگل-ارث به وضوح مکانی تصویر

چند طیفی و وضوح تصویر مورد انتظار پس از بهبود توان تفکیک مکانی تبدیل می‌شود. شکل (۳-۱۰) روند شماتیک این اقدام را نشان می‌دهد.



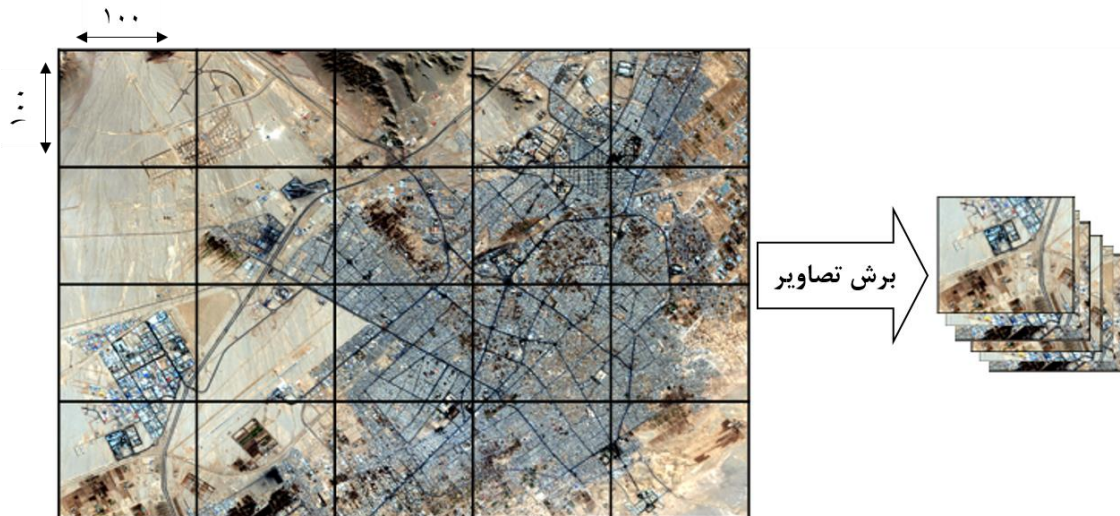
شکل ۳-۱۰: روند شماتیک برش و باز نمونه برداری کاهشی به منظور آماده سازی داده های آموزشی از تصاویر گوگل-ارث

با توجه روند اتخاذ شده در تولید تصاویر آموزشی شکل (۳-۱۰)، تصاویر مورد استفاده به صورت ذاتی هم مرجع بوده و اقدامی مبنی بر هم مرجع سازی آنها نیاز نیست. این تصاویر به برش های  $100 \times 100$  در نسخه هم وضوح با تصویر چند طیفی و  $400 \times 400$  در نسخه ارتقاء یافته مکانی، برای آموزش عمیق شبکه عصبی معرفی شده در بخش (۳-۳-۳) استفاده شده اند. برای تولید این تصاویر آموزشی، پنجره هایی با ابعاد ذکر شده توأم با همپوشانی ۷۵ درصدی از تصاویر گوگل (در نسخه های تنظیم شده از منظر مکانی) استخراج می شود. وجود همپوشانی در نمونه های آموزشی ضامن تأمین نمونه های آموزشی متنوع و بیشتر در روند آموزش تقویتی خواهد بود.



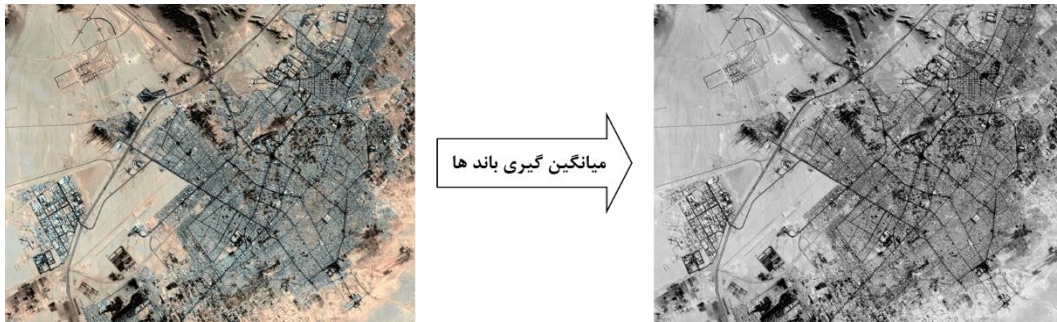
### ۳-۵- تولید تصویر پانکروماتیک

بنا بر روند نمای ارائه شده در شکل (۳-۱)، پس از آموزش شبکه عمیق معرفی شده در بخش (۳-۳-۳) به کمک تصاویر آموزشی گوگل-ارث و ذخیره سازی وزن ها، گام بعد ارتقا وضوح مکانی تصاویر چند طیفی است. در این مرحله باندهای RGB تصاویر چند طیفی سنجنش از دوری به صورت موزائیکی در ابعاد  $100 \times 100$  برش داده می شود (شکل (۳-۱۱)).



شکل ۳-۱۱: روند شماتیک برش موزائیکی تصاویر جهت ورود به شبکه عمیق.

سپس این برش ها وارد شبکه آموزش دیده شده و خروجی برش هایی با ترکیب باند RGB و وضوح چهار برابری نسبت به تصاویر وارد شده به شبکه خواهند بود. در گام بعدی تصاویر بریده شده به هم چسبانده می شوند و به شکل یک تصویر واحد از منطقه مورد نظر به دست خواهد آمد. با میانگین گیری از باندهای تصویر تولید شده یک تصویر تک بانده به دست خواهد آمد که این تصویر به عنوان تصویر پانکروماتیک وارد مرحله بعدی می شود (شکل (۳-۱۲)).



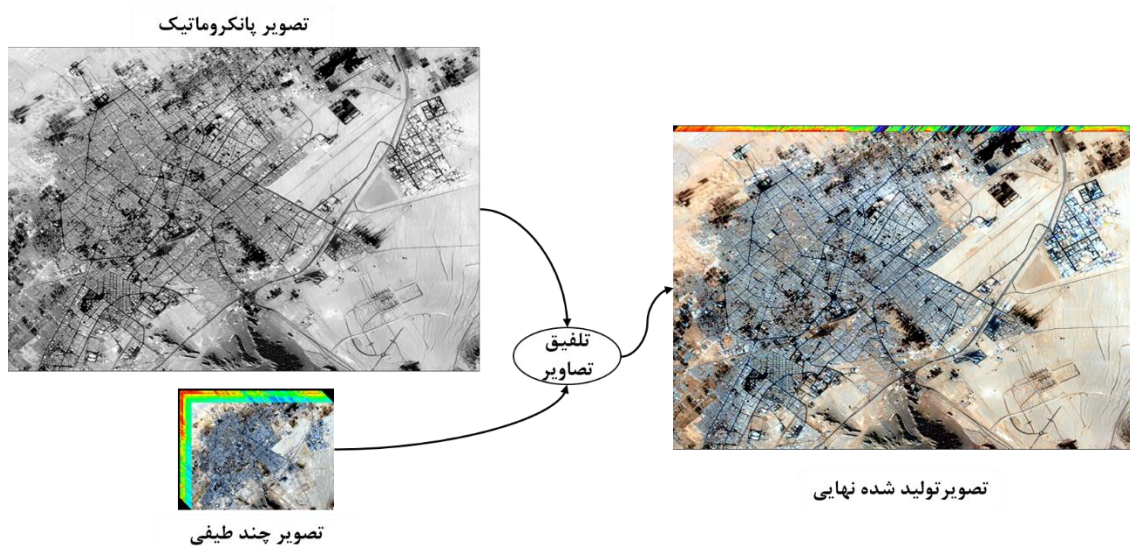
شکل ۳-۱۲: تولید تصویر پانکروماتیک

### ۳-۶- تلفیق تصاویر

تلفیق تصویر یک رویکرد مرسوم در ایجاد هم‌افزایی از مزایای تصاویر اخذشده توسط سنجنده‌های مختلف است. یکی از این مزیت‌ها بهره‌مندی هم‌زمان از جزئیات مکانی و محتوای طیفی تصاویر اخذشده از تصاویر پانکروماتیک و چند طیفی است. مطابق با روند نمای شکل (۳-۱)، پس از تولید تصاویر با وضوح مکانی بالاتر از تصاویر چند طیفی، این تصاویر با تصاویر چند طیفی تلفیق می‌شوند. در این مرحله از روش گرام اشمیت به‌عنوان روش تلفیق استفاده شده است.

تبدیل گرام اشمیت روشی در جبر خطی برای متعامد سازی مجموعه‌ای از بردارها است. اولین بار توسط کداک از این تبدیل برای تلفیق تصاویر استفاده نمود [۷۵]. در این روش، یک تصویر پانکروماتیک هم‌اندازه با تصویر چند طیفی از طریق میانگین‌گیری وزن‌دار از باندهای طیفی شبیه‌سازی شده و سایر باندهای طیفی نسبت به آن و نسبت به یکدیگر از طریق تبدیل گرام-اشمیت متعامد سازی می‌شوند. سپس باند پانکروماتیک اصلی جایگزین باند شبیه‌سازی شده و به‌تبع آن سایر باندهای طیفی به‌صورت متعامد بر آن به محاسبه می‌شوند. در انتها با اعمال تبدیل معکوس گرام-اشمیت، باندهای طیفی در ابعادی هم‌اندازه با تصویر پانکروماتیک بازسازی می‌شوند. ضرایب مربوط به بازسازی باندهای طیفی متعامد در یک‌روند تکراری و به‌ازای هر تصویر برآورد می‌شود. جزئیات مربوط به اجرای

این روش تلفیق در [۷۵] قابل دسترس است. امکان به‌کارگیری این روش در تصاویر چند و ابر طیفی علاوه بر نتایج مطلوب گزارش شده برای این روش [۷۵] از مزیت‌های آن شمرده می‌شود (شکل ۳-۱۳).

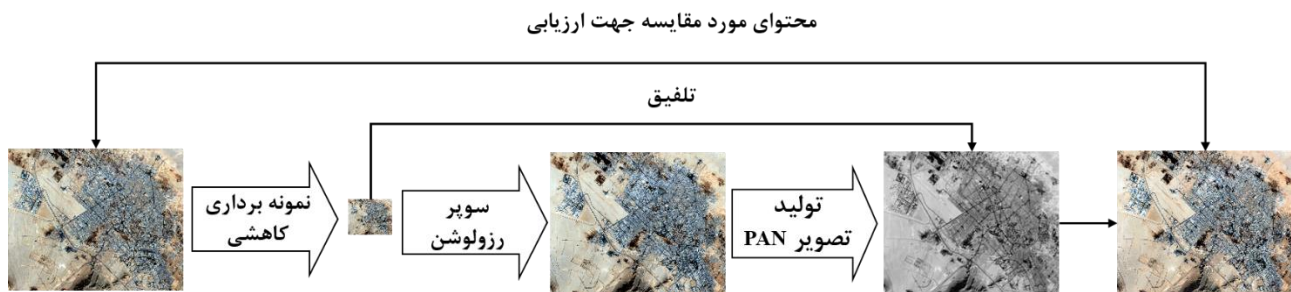


شکل ۳-۱۳: تلفیق تصویر پانکروماتیک تولیدشده و تصویر چند طیفی.

### ۳-۷- رویکردهای ارزیابی نتایج

متناسب با سازوکار روش پیشنهادشده در این تحقیق، رویکردهای ارزیابی نتایج با تفاوت‌های اندکی شبیه به رویکردهای مرسوم در ارزیابی نتایج تلفیق تصاویر است [۳۸]. محصول نهایی در روش پیشنهادی یک تصویر چند طیفی با توان تفکیک مکانی بالاتر از نسخه اولیه است. چنین محصولاتی در دو سطح میزان حفظ (۱ جزئیات مکانی و ۲ محتوای طیفی مورد ارزیابی قرار می‌گیرند. در روند ارزیابی میزان حفظ محتوای مکانی، یک تصویر پانکروماتیک با توان تفکیک مکانی متناظر با تصویر ارتقاءیافته مبنای مقایسه قرار گرفته است. تصاویر ماهواره‌ای مورد استفاده در این تحقیق لندست-۸ و سنتینل-۲ بوده که تصاویر پانکروماتیک سنجنده کارتوست (P5) با توان تفکیک مکانی ۲/۵ متری برای ارزیابی این بخش استفاده شده است.

در روند ارزیابی میزان حفظ محتوای طیفی نیز به دلیل عدم وجود تصاویر با توان تفکیک طیفی و مکانی مشابه همانند رویکردهای رایج ارزیابی نتایج تلفیق در شرایط مشابه دنبال شده است. در این رویکردها تصویر چند طیفی از طریق نمونه برداری کاهشی به تصویری با توان تفکیک مکانی پایین تر تبدیل شده و سپس روند بازسازی نسخه با وضوح بالاتر از طریق روش پیشنهادی به اجرا می رسد. نتایج تولید شده تصاویری با توان تفکیک طیفی و مکانی مشابه با تصاویر چند طیفی اولیه بوده که امکان مقایسه میزان اعوجاجات طیفی در آن ها وجود خواهد داشت. شکل (۳-۱۴) این رویکرد را به صورت شماتیک نشان می دهد.



شکل ۳-۱۴: روند شماتیک تولید محتوا مورد مقایسه جهت ارزیابی میزان حفظ محتوای طیفی در روش پیشنهادی

شاخص های متنوعی به منظور ارزیابی نتایج تلفیق تصاویر وجود داشته که به دودسته معیارهای طیفی و مکانی قابل تفکیک هستند. در این تحقیق از شاخص های (۱) حداکثر نسبت سیگنال به نویز (PSNR<sup>۱</sup>)، (۲) شباهت ساختاری (SSIM<sup>۲</sup>) و (۳) ضریب همبستگی (CC<sup>۳</sup>) به عنوان شاخص های سنجش حفظ اطلاعات طیفی و شاخص های (۱) ضریب همبستگی مکانی (SCC)، (۲) کامل بودن لبه های تصویری (Cm<sup>۴</sup>) و (۳) صحت حفظ محتوای لبه (Cr<sup>۵</sup>) به عنوان رویکرد ارزیابی کفایت مکانی نتایج استفاده شده است که در ادامه به معرفی شاخص های مذکور

1 Peak Signal-to-Noise Ratio  
 2 Structural Self-Interaction Matrix  
 3 Cross Correlation  
 4 Completeness  
 5 Correctness

پردازش شده است. دو مورد ۵ و ۶ که از شاخص‌های حفظ محتوای مکانی هستند، از طریق مقایسه لبه‌های مستخرج بین تصویر تولیدشده و تصویر مبنا در قالب یک ماتریس درهم‌ریختگی دودویی تولید می‌شوند [۳۸, ۷۶, ۷۷].

### ۳-۷-۱- معیارهای طیفی

این معیارها میزان شباهت اطلاعات طیفی دو تصویر در باندهای مختلف را مورد ارزیابی قرار می‌دهند.

### ۳-۷-۱-۱- حداکثر نسبت سیگنال به نویز (PSNR)

این معیار به ازای هر باند طیفی قابل محاسبه بوده و بزرگ بودن آن نشان از حفظ بیشتر محتوای اطلاعاتی در تصویر تلفیق شده خواهد بود. هرچه دو تصویر به هم شبیه‌تر باشند مقدار این شاخص بیشتر می‌شود. اگر دو تصویر کاملاً یکسان باشند رابطه (۳-۴)، به دلیل صفر شدن مخرج، مقدار این شاخص  $\infty$  خواهد شد [۷۸].

$$PSNR = 10 \log_{10} \left[ \frac{1}{\frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N (I_r(i,j) - I_f(i,j))^2} \right] \quad \text{رابطه (۳-۴)}$$

در رابطه (۳-۴)  $M$  و  $N$  ابعاد تصویر،  $I_r$  تصویر مرجع و  $I_f$  تصویر تولیدشده است [۷۹].

### ۳-۷-۱-۲- شباهت ساختاری (SSIM)

شاخص شباهت ساختاری یک شاخص مبتنی بر روشی و کنتراست تصویر بوده که الهام گرفته شده از توانایی‌های درک بصری انسان در مقایسه شباهت دو تصویر است. در محاسبه این شاخص از میانگین و انحراف استاندارد تصویر مرجع و تصویر به دست آمده استفاده می‌شود. بزرگی این شاخص حاکی از شباهت ساختاری بیشتر دو تصویر می‌باشد. اگر دو تصویر کاملاً یکسان باشند مقدار این شاخص برابر ۱ می‌شود. این معیار مقاومت بیشتری در برابر تبدیلات خطی نسبت به PSNR دارد. برای محاسبه معیار SSIM از رابطه (۳-۵) استفاده شده که در آن،  $\mu$  میانگین تصویر مرجع،  $\mu_f$  میانگین تصویر تولیدشده،  $\sigma_r^2$  واریانس تصویر مرجع،  $\sigma_f^2$  واریانس تصویر تولیدشده،  $\sigma_{xy}$  کوواریانس

تصویر مرجع و تولیدشده می باشد. همچنین در رابطه (۳-۵) خواهیم داشت [۳۸].  $k_1 = 0.01$ ،  $k_2 = 0.03$  و  $L = 255$

$$\begin{cases} SSIM_{(r,f)} = \frac{(2\mu_r\mu_f - C_1) \times (2\sigma_{rf} + C_2)}{(\mu_r^2 + \mu_f^2 - C_1) \times (\sigma_r^2 + \sigma_f^2 + C_2)} \\ C_1 = (K_1 L)^2, C_2 = (K_2 L)^2 \end{cases} \quad \text{رابطه (۳-۵)}$$

### ۳-۷-۱-۳- ضریب همبستگی (CC)

این معیار شباهت تصویر مرجع با تصویر تولیدشده را مشخص می کند. هرچه مقدار عددی به ۱ نزدیک تر باشد بیانگر شباهت بیشتر اطلاعات طیفی تصاویر است. در رابطه (۳-۶)  $\mu_r$  میانگین تصویر مرجع،  $\mu_f$  میانگین تصویر تولیدشده،  $f$  تصویر مرجع و  $r$  تصویر تولیدشده می باشد [۷۶].

$$CC = \frac{\sum_{x=1}^m \sum_{y=1}^n (f(x,y) - \mu_f)(r(x,y) - \mu_r)^2}{\sqrt{\sum_{x=1}^m \sum_{y=1}^n (f(x,y) - \mu_f)^2 (r(x,y) - \mu_r)^2}} \quad \text{رابطه (۳-۶)}$$

### ۳-۷-۲- معیارهای هندسی

این معیارها شباهت اطلاعات هندسی دو تصویر مانند لبه ها و خصوصیات ساختاری را مورد ارزیابی قرار می دهند.

### ۳-۷-۲-۱- ضریب همبستگی مکانی (SCC)

هرچه همبستگی مؤلفه های بالا بین تصویر مرجع و تصویر تولیدشده بیشتر باشد، بیانگر صحت اطلاعات مکانی تصویر تولیدشده می باشد. پنجره بالا گذاری که در این تحقیق مورد استفاده قرار گرفته است به شرح زیر می باشد [۷۶]. رابطه (۳-۷)

$$Mask = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad \text{رابطه (۳-۷)}$$

به کمک پنجره بالا گذر معرفی شده لبه ها را در دو تصویر مرجع و تولیدشده، استخراج نموده و سپس از معیار ضریب همبستگی (۳-۷-۲-۱) استفاده می شود.

### ۳-۷-۲-۲- کامل بودن لبه‌های تصویری (Cm) و صحت حفظ محتوای لبه (Cr)

در معیار های کامل بودن و صحت داشتن که دو مورد از معیار های هندسی ارزیابی تصاویر هستند، ابتدا باید ۳ مولفه محاسبه شوند که این ۳ مولفه عبارتند از: (۱) TP معرف تعداد پیکسل های شناسایی شده صحیح<sup>۱</sup>، (۲) FN معرف تعداد پیکسل های شناسایی شده غیر صحیح<sup>۲</sup> و (۳) FP معرف تعداد پیکسل های شناسایی نشده<sup>۳</sup>. پارامتر های ذکر شده از طریق مقایسه لبه‌های مستخرج بین تصویر تولید شده و تصویر مبنا در قالب یک ماتریس درهم‌ریختگی دو دویی تولید می‌شوند و سپس به کمک روابط (۳-۸ و ۳-۹) میزان کامل بودن لبه های تصویری و صحت محتوای لبه ارزیابی می‌شود [۷۷].

$$Cm = \frac{TP}{TP+FN} \quad \text{رابطه (۳-۸)}$$

$$Cr = \frac{TP}{TP+FP} \quad \text{رابطه (۳-۹)}$$

### ۳-۸- جمع بندی

در این فصل تمام مراحل روش پیشنهادی جهت بهبود کیفیت مکانی تصاویر ماهواره ای با وضوح مکانی متوسط از طریق روش های یادگیری عمیق به صورت کامل و با جزئیات توضیح داده شد. به طوری که در ابتدا ساختار کلی شبکه عمیق GAN و معماری های استفاده شده در آن و سپس شبکه عمیق GAN پیشنهادی با جزئیات تمامی لایه های استفاده شده مورد بررسی قرار گرفت (بخش ۳-۳) و پس از آن روندی برای پیش پردازش و آماده سازی داده های آموزشی ارائه شد (بخش ۳-۴). در بخش های (۳-۵) و (۳-۶) روند تولید تصویر پانکروماتیک و تلفیق تصویر پانکروماتیک تولید شده و تصویر چند طیفی سنجنده سنجش از دوری به طور کامل بررسی شد و در نهایت،

---

<sup>۱</sup> True Positive

<sup>۲</sup> False Negative

<sup>۳</sup> False Positive

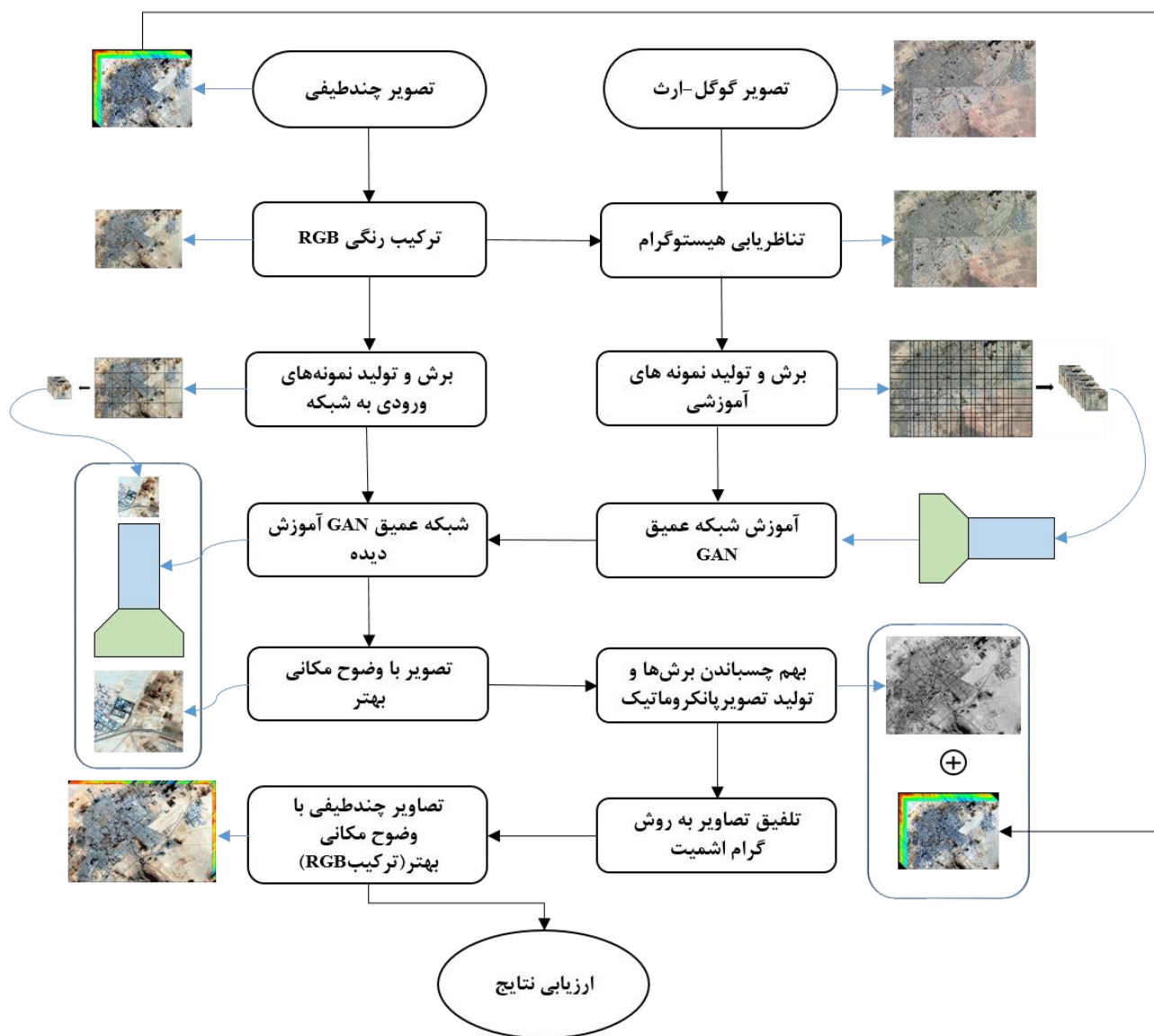
معیارهای ارزیابی نتایج در دو زیر گروه ارزیابی هندسی و ارزیابی طیفی به صورت کامل شرح داده شد (بخش ۳-۷).



## فصل ۴: پیاده‌سازی و ارزیابی نتایج

### ۴-۱- مقدمه

در این فصل از تحقیق، به پیاده‌سازی و ارزیابی روش پیشنهاد شده جهت بهبود کیفیت مکانی تصاویر ماهواره ای با وضوح مکانی متوسط از طریق روش‌های یادگیری عمیق پرداخته شده است. در این راستا، ابتدا مناطق مورد مطالعه و داده‌های استفاده شده به‌طور کامل تشریح شده‌اند (بخش ۴-۲). سپس، در بخش (۴-۳) روش پیشنهادی طبق روند نمایی ارائه شده شکل (۴-۱) پیاده‌سازی شده و به کمک معیارهای ارزیابی ارائه شده در بخش (۴-۷) مورد ارزیابی قرار می‌گیرد.



شکل ۴-۱: روند نمای پیاده سازی روش پیشنهادی

## ۴-۲- مناطق مطالعاتی و داده های مورد استفاده

در این تحقیق از چهار مجموعه داده جهت پیاده سازی و ارزیابی روش پیشنهادی استفاده شده است. (۱) باند پانکروماتیک سنجنده کارتوست، (۲) تصاویر گوگل-ارث (<https://earth.google.com/web>)، (۳) باندهایی از تصاویر متعلق به سنجنده لندست-۸ (<https://earthexplorer.usgs.gov>) و (۴) باندهایی از تصاویر متعلق به

سنجنده سنتینل-۲ (<https://scihub.copernicus.eu>). تصاویر ماهواره ای سنجنده لندست-۸ (جدول (۴-۱)) متشکل از ۱۱ باند مختلف در محدوده نور مرئی، مادون قرمز نزدیک، موج کوتاه مادون قرمز و طیف مادون قرمز حرارتی با توان تفکیک مکانی ۱۵، ۳۰ و ۱۰۰ متری می‌باشد که در این تحقیق باندهای ۳۰ و ۱۵ متری این سنجنده مورد استفاده قرار گرفته است و به دلیل اختلاف زیاد توان تفکیک مکانی باندهای ۱۰۰ متری نسبت به محصول تولیدی با وضوح مکانی بالا از آنها استفاده نشده است.

جدول ۱-۴: معرفی باندهای طیفی سنجنده لندست-۸

توان تفکیک مکانی (m)	میانگین طول موج (nm)	باندها
۳۰ متر	۴۴۳	باند ساحلی <sup>۱</sup>
۳۰ متر	۴۸۲	باند آبی
۳۰ متر	۵۶۲	باند سبز
۳۰ متر	۶۵۵	باند قرمز
۳۰ متر	۸۶۵	باند مادون قرمز نزدیک
۳۰ متر	۱۶۱۰	باند مادون قرمز طول موج کوتاه
۳۰ متر	۲۲۰۰	باند مادون قرمز طول موج کوتاه
۱۵ متر	۵۹۰	باند پانکروماتیک
۳۰ متر	۱۳۷۵	باند سیروس <sup>۲</sup>
۱۰۰ متر	۱۰۸۰	باند مادون قرمز طول موج بلند
۱۰۰ متر	۱۲۰۰	باند مادون قرمز طول موج بلند

Coastal<sup>۱</sup>

Cirrus<sup>۲</sup>

تصاویر ماهواره ای سنجنده سنتینل-۲ نیز متشکل از ۱۳ باند (جدول (۲-۴)) مختلف در محدوده مرئی و فروسرخ با توان تفکیک مکانی ۱۰، ۲۰ و ۶۰ متری می‌باشند که در این تحقیق باند های ۱۰ و ۲۰ متری این سنجنده مورد استفاده قرار گرفته است و در این سنجنده نیز همانند قبل از باند های ۶۰ متری استفاده نشده است.

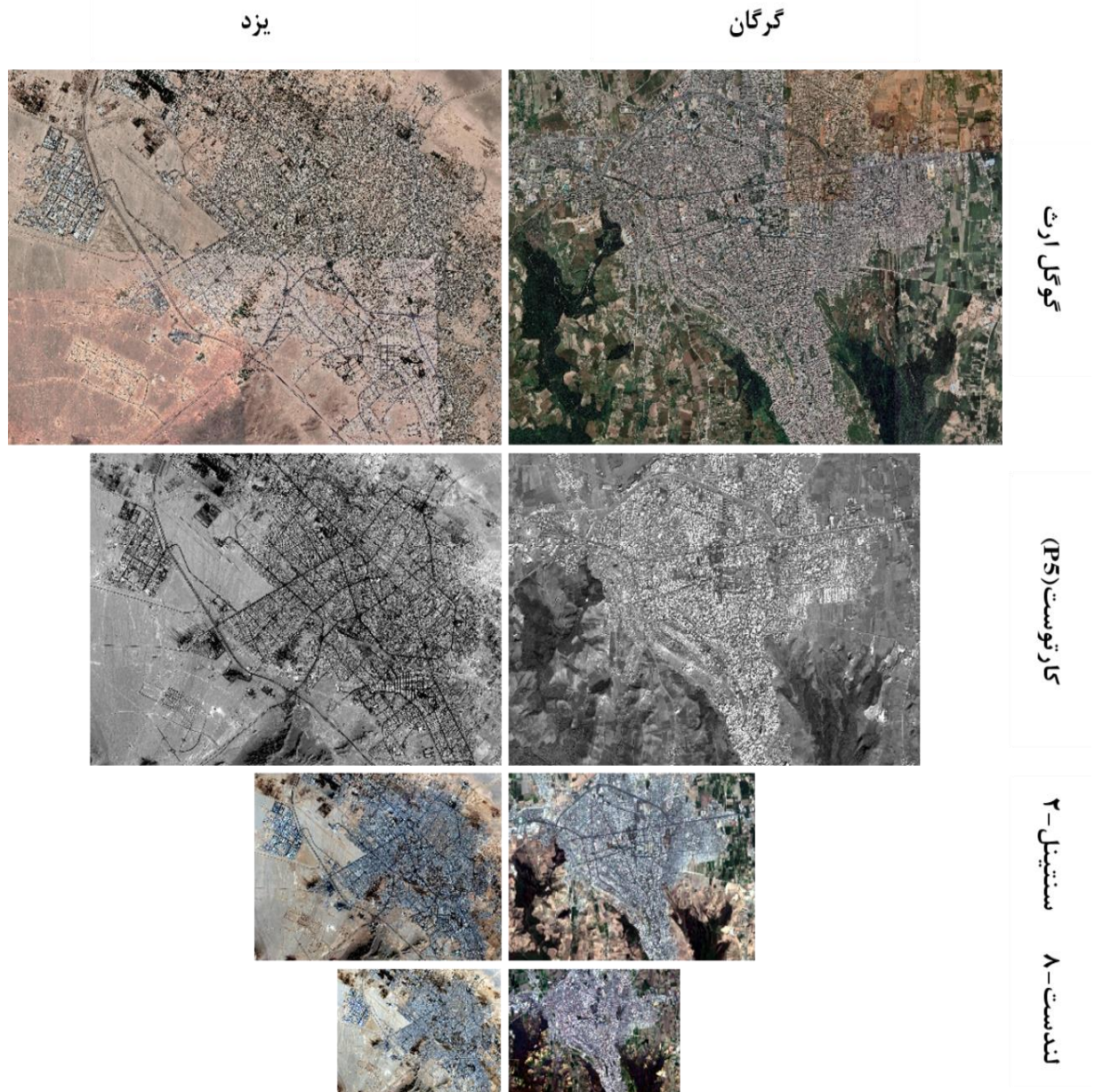
جدول ۲-۴: معرفی باند های طیفی سنجنده سنتینل-۲

باند ها	میانگین طول موج (nm)	توان تفکیک مکانی (m)
باند ساحلی	۴۴۳	۶۰ متر
باند آبی	۴۹۰	۱۰ متر
باند سبز	۵۶۰	۱۰ متر
باند قرمز	۶۶۵	۱۰ متر
باند مادون قرمز مرئی و نزدیک <sup>۱</sup> (VNIR)	۷۰۵	۲۰ متر
باند مادون قرمز مرئی و نزدیک	۷۴۰	۲۰ متر
باند مادون قرمز مرئی و نزدیک	۷۸۳	۱۰ متر
باند مادون قرمز مرئی و نزدیک	۸۴۲	۲۰ متر
باند مادون قرمز مرئی و نزدیک	۸۶۵	۲۰ متر
باند موج کوتاه مادون قرمز (SWIR)	۹۴۰	۶۰ متر
باند موج کوتاه مادون قرمز	۱۳۷۵	۶۰ متر
باند موج کوتاه مادون قرمز	۱۶۱۰	۲۰ متر
باند موج کوتاه مادون قرمز	۲۱۹۰	۲۰ متر

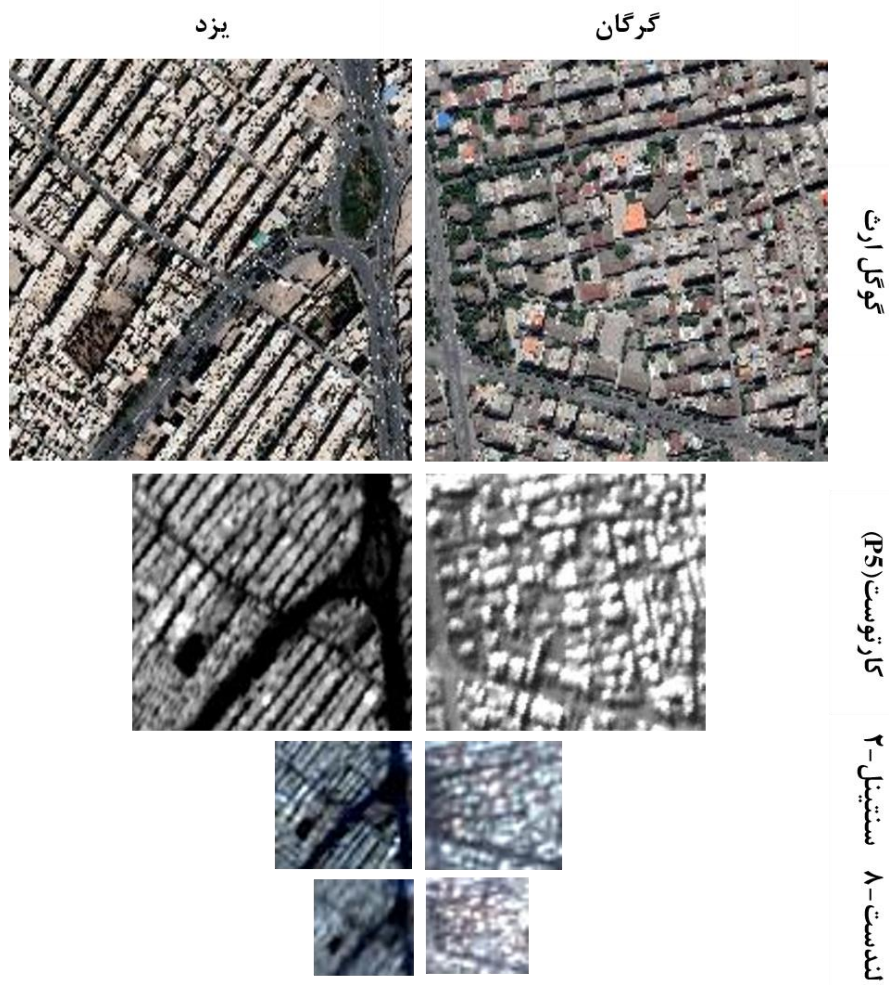
از باند پانکروماتیک ماهواره کارتوست با وضوح مکانی ۲/۵ متر، به منظور تصویر مرجع در ارزیابی مکانی و از تصاویر گوگل-ارث در راستای آموزش شبکه عمیق استفاده شده اند. منطقه مورد مطالعه در این تحقیق شهر های یزد و

<sup>۱</sup>Visible and Near Infrared

گرگان هستند که مواردی همچون در دسترس بودن، وجود تنوع در پوشش زمین و بافت شهری متمایز دلایل انتخاب این دو مجموعه داده بوده است (شکل ۴-۲ و ۴-۳). موقعیت و وسعت این مناطق در جدول (۴-۳) ارائه شده اند.



شکل ۴-۲: تصاویر مورد استفاده از شهر یزد و گرگان



شکل ۴-۳: شمای نزدیک تصاویر سنجنده های مختلف از شهر یزد و گرجان.

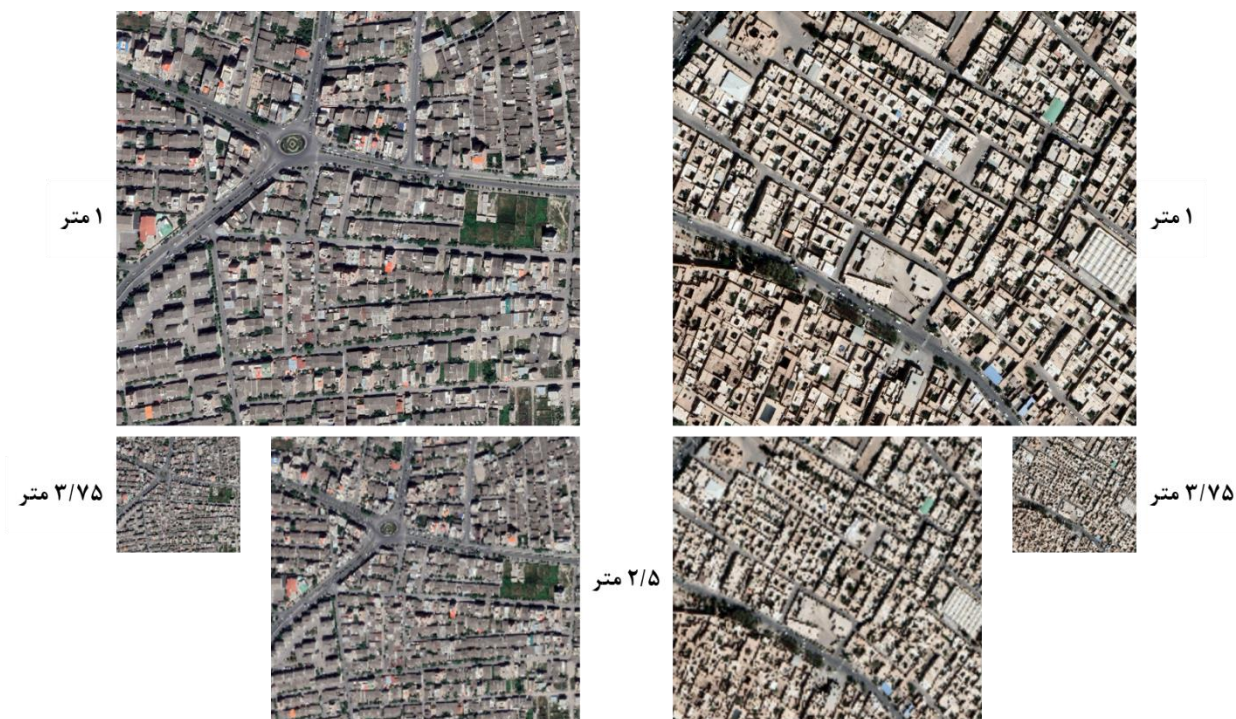
جدول ۳-۴: جزئیات مربوط به منطقه مورد مطالعه در ارزیابی نتایج روش پیشنهادی

شهر	موقعیت جغرافیایی	وسعت (km <sup>2</sup> )
گرجان	$\varphi = 36^{\circ}49'$ $\lambda = 54^{\circ}26'$	۳۲۶
یزد	$\varphi = 31^{\circ}53'$ $\lambda = 54^{\circ}20'$	۸۸۳

## ۴-۳- پیاده سازی و ارزیابی نتایج

در این بخش پیاده سازی روش پیشنهادی و نتایج ارزیابی آن گزارش شده است. با توجه به عدم یکنواختی توان تفکیک مکانی باندهای طیفی تصاویر چندطیفی سنتینل-۲ و لندست-۸، ابتدا از طریق روش تلفیق گرام-اشمیت توان تفکیک باندهای طیفی این تصاویر همگون شده اند. به عبارت بهتر، باندهای طیفی ۲۰ متری سنجنده سنتینل-۲ با میانگین باندهای طیفی ۱۰ متری و باندهای طیفی ۳۰ متری سنجنده لندست-۸ با باند پانکروماتیک ۱۵ متری تلفیق شده و توان تفکیک مکانی آنها یکسان شده است. عدم استفاده از باندهای طیفی ۶۰ متری در سنجنده سنتینل-۲ و ۱۰۰ متری در سنجنده لندست-۸ به دلیل کیفیت پایین فرایند تلفیق برای این باندهای طیفی در نتایج تولید شده بود. به این ترتیب تمامی باندهای استفاده شده از سنجنده سنتینل-۲ در این پژوهش با وضوح مکانی ۱۰ متر و تمامی باندهای استفاده شده از سنجنده لندست-۸ با وضوح مکانی ۱۵ متر هستند.

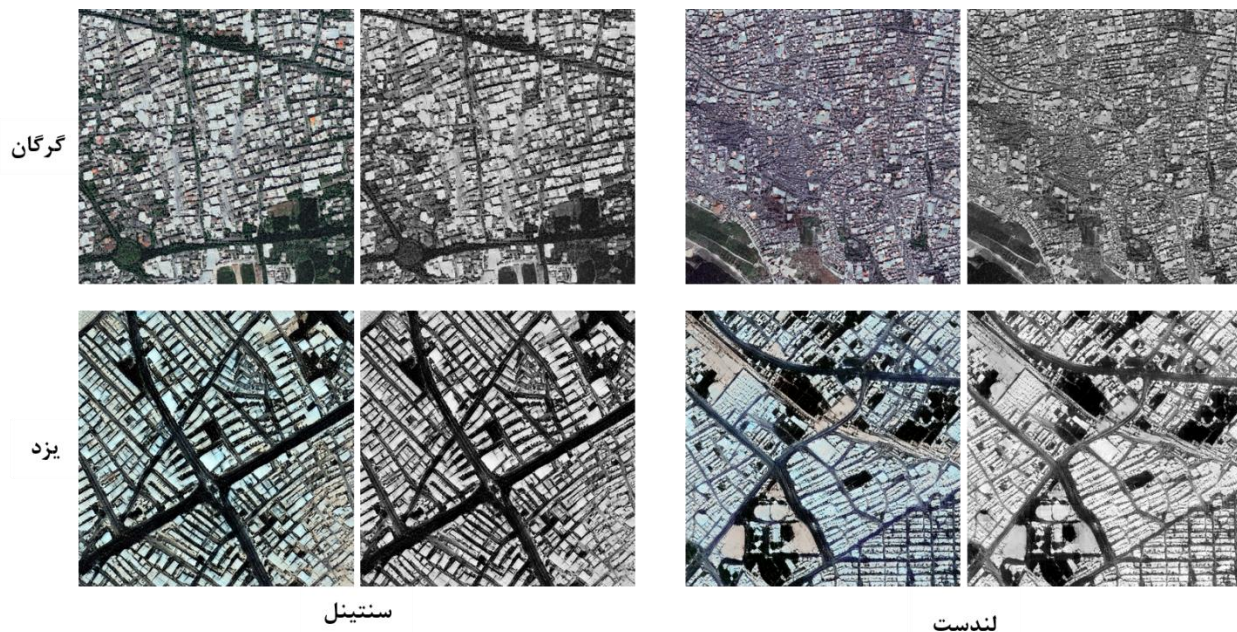
بطور موازی، تصاویر مستخرج از سامانه گوگل-ارث از طریق نمونه برداری کاهشی به دو مجموعه آموزشی برای تصاویر سنتینل-۲ و لندست-۸ تبدیل شده اند. قبل از این اقدام، هیستوگرام این تصاویر از طریق تناظریابی هیستوگرام به ترکیب رنگی مرئی از تصاویر سنتینل-۲ و لندست-۸ شبیه شده است. با فرض توان تفکیک مکانی یک متری برای تصاویر مستخرج از سامانه گوگل ارث، تصاویر باز نمونه برداری شده برای آموزش تقویتی شبکه عصبی عمیق در مجموعه اول به وضوح مکانی ۲/۵ و ۱۰ متری تبدیل یافته و در مجموعه دوم این فرایند منجر به تولید تصاویری با وضوح ۳/۷۵ و ۱۵ متری شده است. از مجموعه اول جهت آموزش تقویتی شبکه عصبی Real-ESRGAN مربوط به تصاویر سنتینل-۲ و از مجموعه دوم برای آموزش جهت ارتقاء وضوح مکانی تصاویر لندست-۸ استفاده شده است (شکل (۴-۴)).



شکل ۴-۴: نمونه برداری کاهشی تصاویر ۱ متری گوگل ارث برای آموزش شبکه در راستای بهبود وضوح مکانی تصاویر سنجنده لندست-۸ و سنتینل-۲.

در گام سوم از روند پیاده‌سازی، آموزش تقویتی شبکه عصبی Real-ESRGAN در دو نسخه به کمک مجموعه‌های اول و دوم از تصاویر آماده‌سازی شده گوگل-ارث به اجرا رسیده و در ادامه از آنها برای ارتقاء وضوح مکانی تصاویر سنجنده‌های سنتینل-۲ و لندست-۸ استفاده شده است. مطابق با روندنمای ارائه شده در شکل (۴-۱)، از شبکه عصبی Real-ESRGAN صرفاً با هدف تولید محتوای متناظر با تصاویر پانکروماتیک در روند تلفیق تصاویر استفاده شده است. با توجه به اینکه در روند ارتقاء وضوح مکانی صرفاً از باندهای مرئی سنجنده‌های سنتینل-۲ و لندست-۸ استفاده شده و محصول تولید شده نیز یک تصویر RGB است؛ محصول تولید شده از طریق میانگین‌گیری به یک تصویر پانکروماتیک تبدیل می‌گردد. برخی از نتایج این فرآیند در شکل (۴-۵) به نمایش گذاشته شده است.





شکل ۴-۵: تولید تصویر پانکروماتیک از تصاویر چند طیفی تولید شده از شبکه. در جفت تصاویر، تصویر سمت چپ تصویر چند طیفی تولید شده و تصویر سمت راست تصویر پانکروماتیک مستخرج از آن است.

در نهایت، با استفاده از روش تلفیق تصاویر گرام-اشمیت تصویر چندطیفی با وضوح مکانی بهتر تولید شده است. روند آموزش تقویتی و سایر اقدامات پردازش در یک دستگاه رایانه با پردازنده Core i7، 16 گیگابایت RAM و کارت گرافیک NVIDIA GeForce GTX 3060 Ti دارای 6 گیگابایت حافظه GPU<sup>1</sup> به اجرا رسیده است. روش و جزئیات مرتبط با روند آموزش تقویتی شبکه عصبی Real-ESRGAN در جدول (۴-۴) گزارش شده است [۷۱].

جدول ۴-۴: روش و تنظیمات مرتبط با روند آموزش تقویتی شبکه عصبی Real-ESRGAN

روش یادگیری	ADAM
تعداد اپک‌های آموزش	۳۰
نرخ کاهش وزن <sup>۲</sup>	$5 \times 10^{-3}$
نرخ یادگیری	$2 \times 10^{-4}$

<sup>1</sup> Graphics Processing Unit

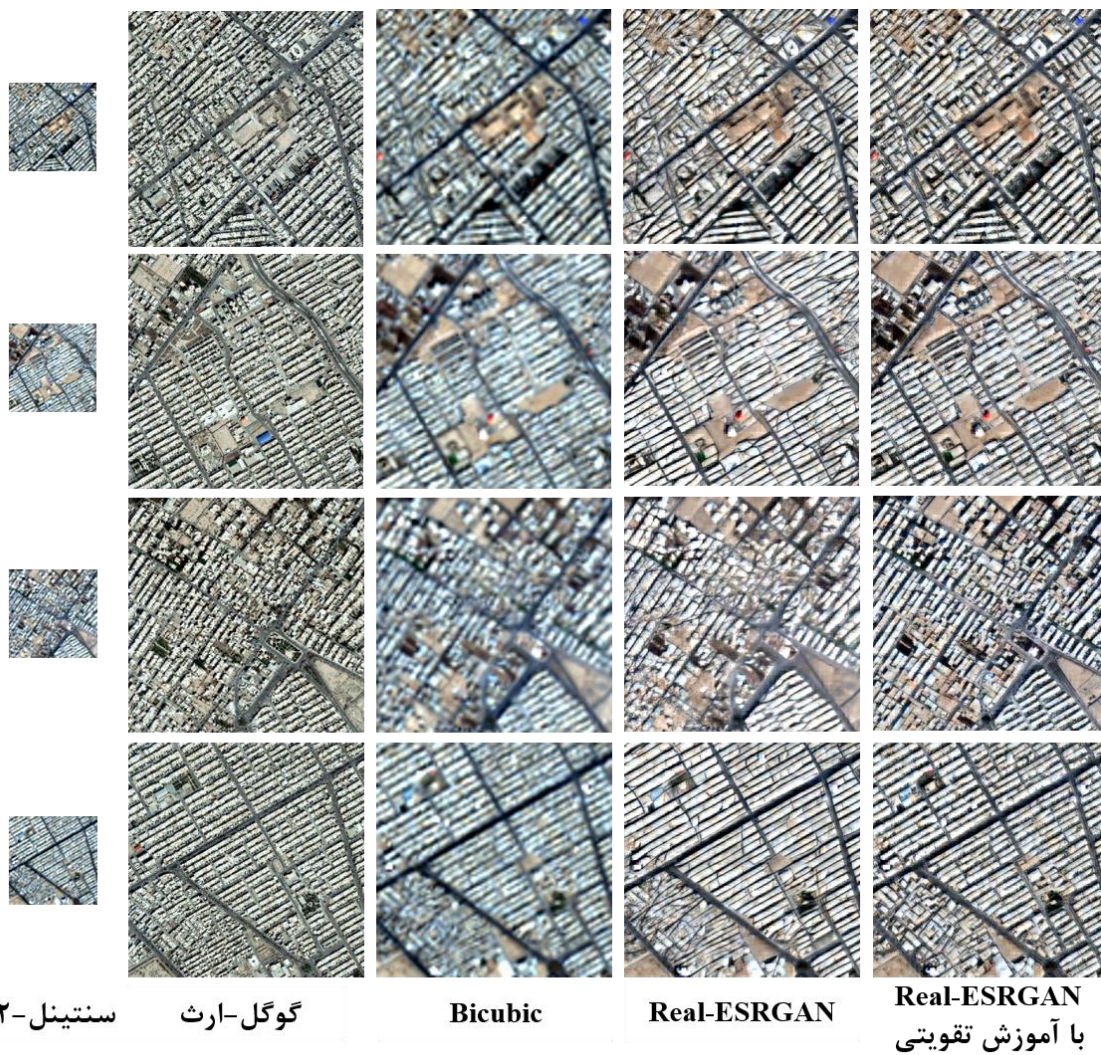
<sup>۲</sup> Weight decay

همچنین تعداد داده آموزشی و زمان تقریبی آموزش برای هر شهر و هر سنجنده در جدول (۴-۵) ارائه شده است.

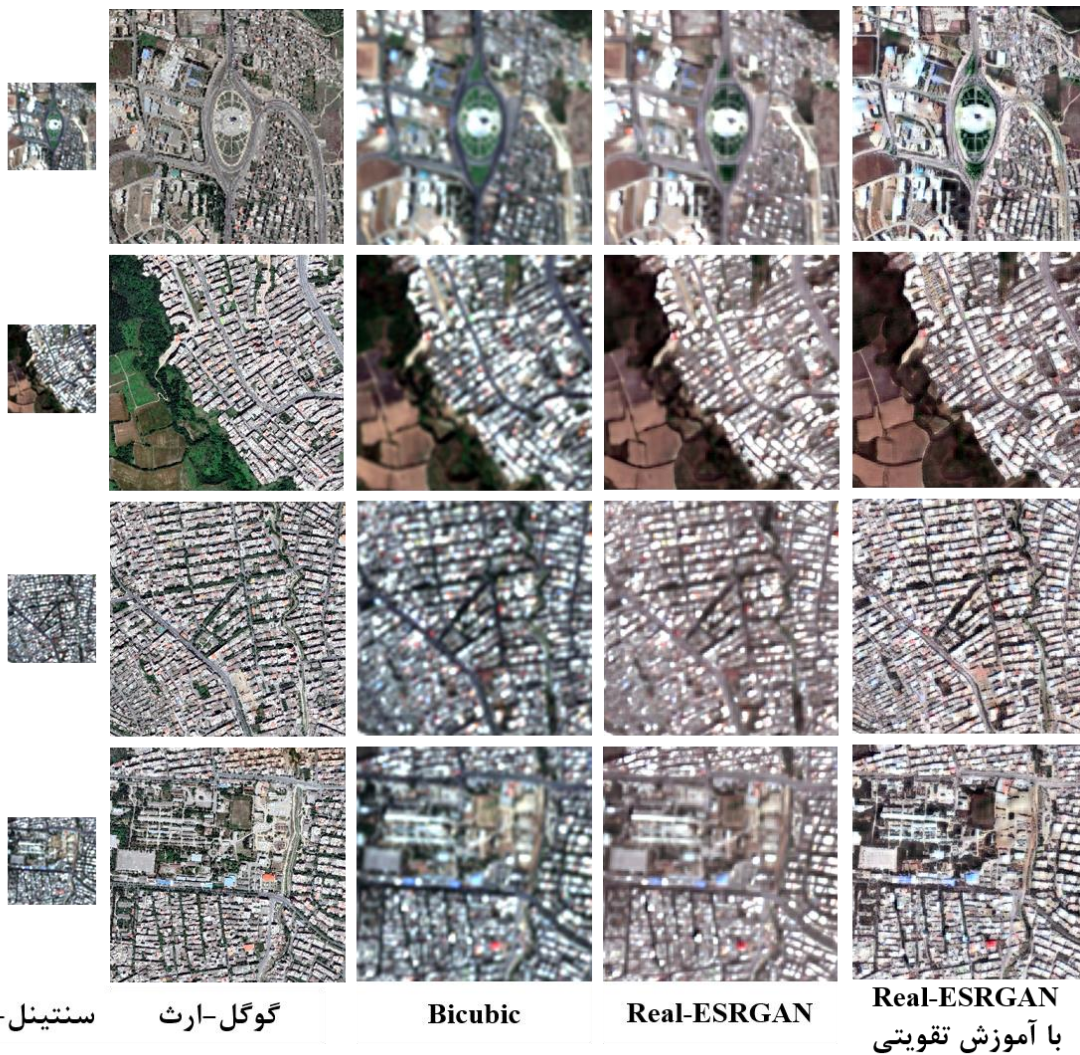
جدول ۴-۵: تعداد داده های آموزشی استفاده شده در این تحقیق.

سنجنده	شهر	تعداد داده آموزشی	زمان تقریبی آموزش
لندست-۸	یزد	۲۷۵۰۰	۴۸ ساعت
لندست-۸	گرگان	۸۳۰۰	
سنتینل-۲	یزد	۴۲۴۰۰	
سنتینل-۲	گرگان	۱۲۹۰۰	

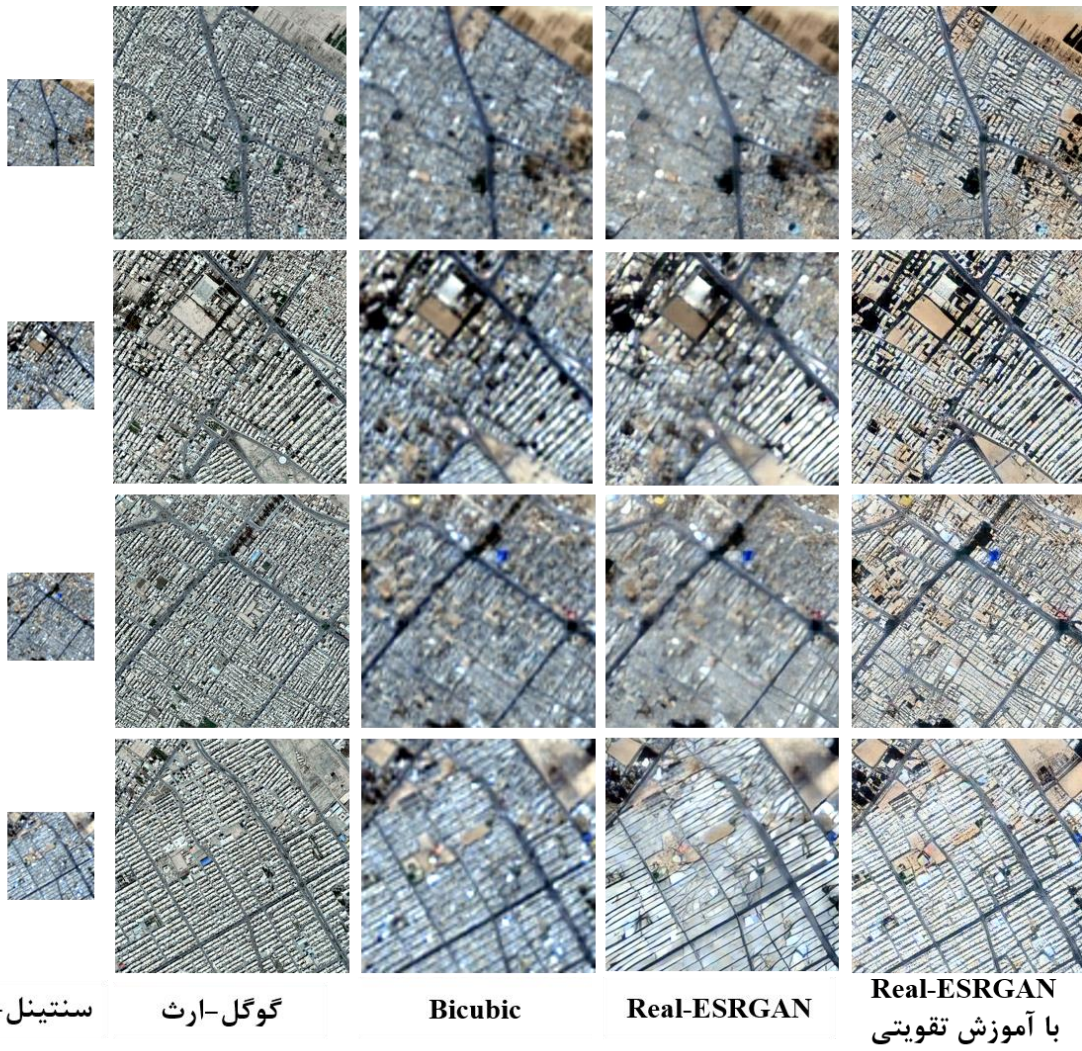
ارزیابی روش پیشنهادی در این تحقیق به کمک دو روش مقایسه ای صورت پذیرفته است. عدم بکلگیری تکنیک آموزش تقویتی و بکلگیری روش درونیابی سنتی bicubic [۸۰] دوراهکار مقایسه ای مورد استفاده هستند. شکل های (۴-۶ و ۴-۷) به ترتیب نتایج کسب شده از روش پیشنهادی برای تصاویر سنتینل-۲ در دو شهر یزد و گرگان و شکل های (۴-۸ و ۴-۹) به ترتیب نتایج کسب شده از روش پیشنهادی برای تصاویر لندست-۸ در دو شهر یزد و گرگان در این تحقیق را نشان می دهد. در این تصاویر صرفاً ترکیب های رنگی مرئی برای نمایش نتایج استفاده شده است.



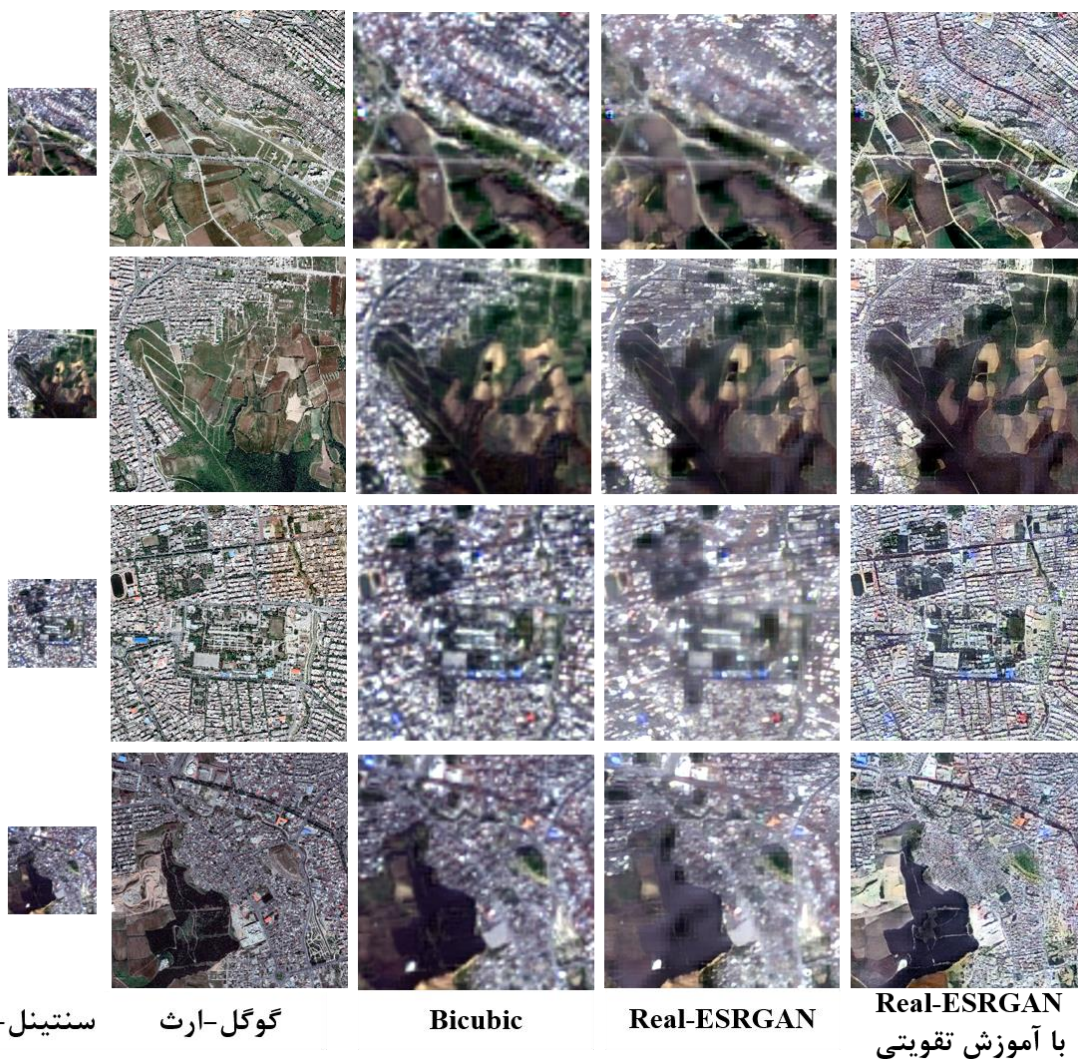
شکل ۴-۶: نمایی از نتایج کسب شده از روش پیشنهادی و سایر روش‌های مقایسه‌ای در مورد تصاویر سنجنده سنتینل-۲؛ شهر یزد.



شکل ۴-۷: نمایی از نتایج کسب شده از روش پیشنهادی و سایر روش‌های مقایسه‌ای در مورد تصاویر سنجنده سنتینل-۲؛ شهر گرگان.



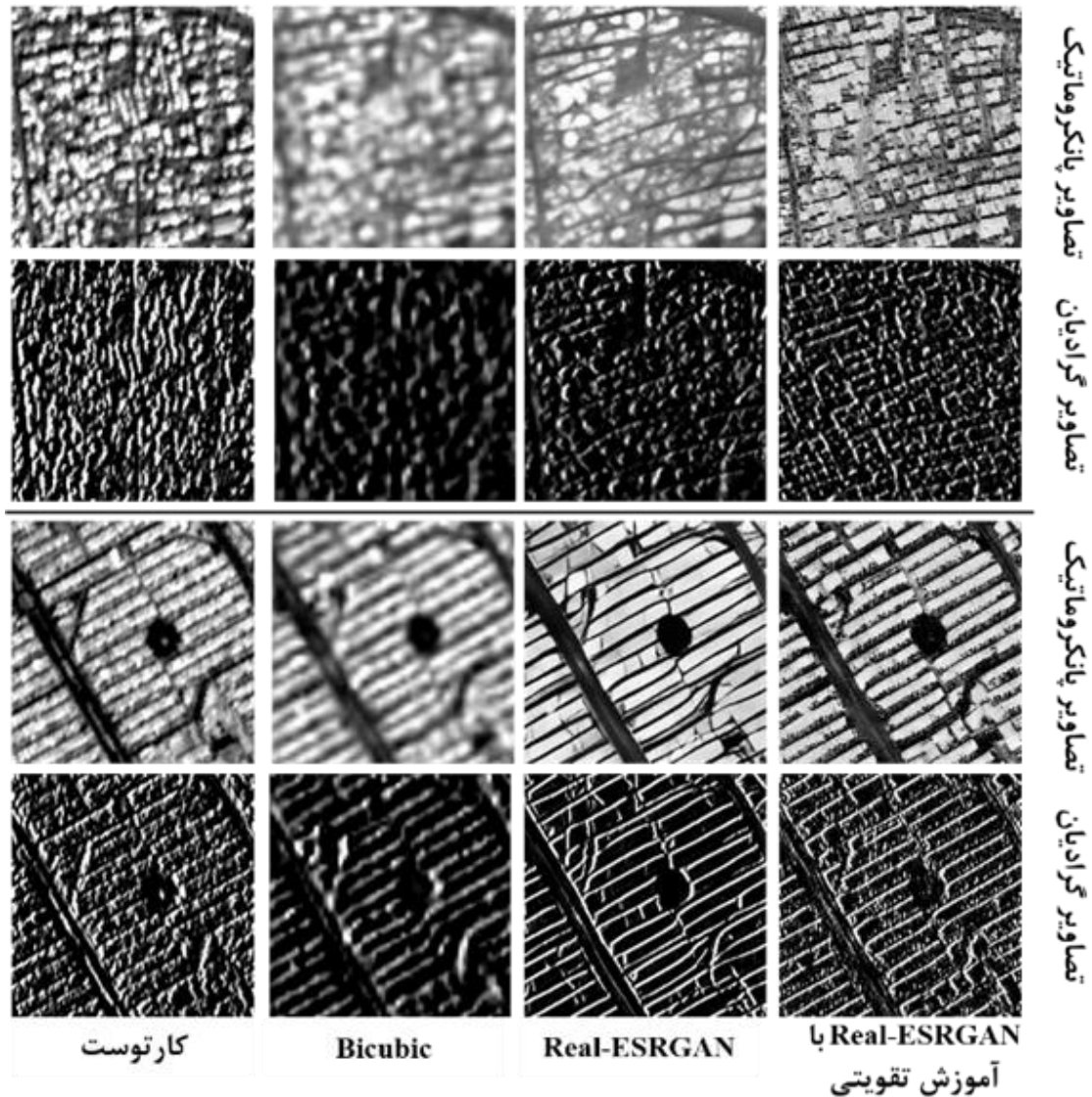
شکل ۴-۸: نمایی از نتایج کسب شده از روش پیشنهادی و سایر روش‌های مقایسه‌ای در مورد تصاویر سنجنده لندست-۸؛ شهر یزد.



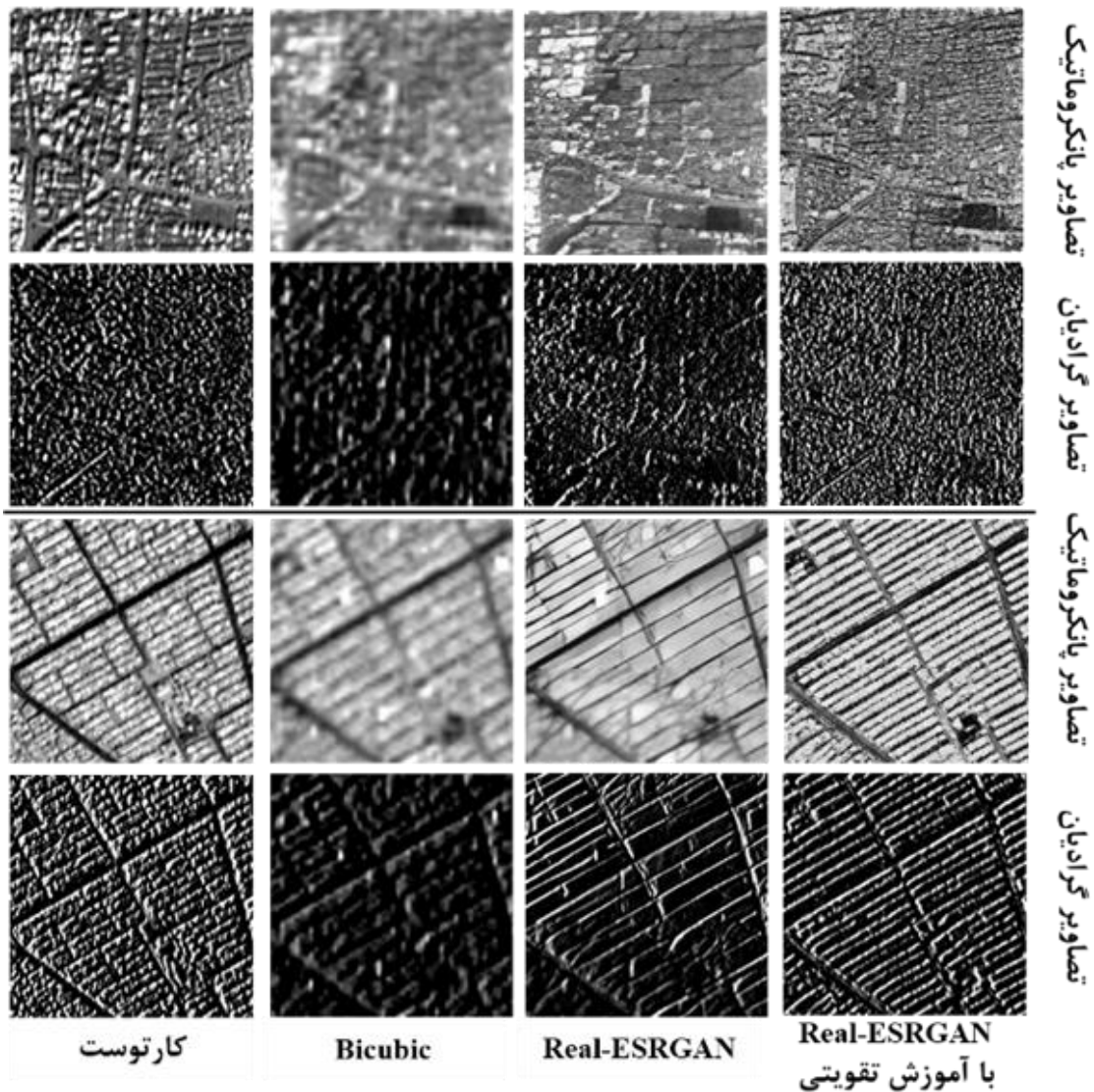
شکل ۴-۹: نمایی از نتایج کسب شده از روش پیشنهادی و سایر روش‌های مقایسه‌ای در مورد تصاویر سنجنده لندست-۸؛ شهر گرگان.

مطابق با نتایج ارائه شده در شکل‌های (۴-۶، ۴-۷، ۴-۸ و ۴-۹)، روش پیشنهادی از نظر بصری نتایج بهتری را نسبت به دوروش مقایسه‌ای کسب نموده و جزئیات مکانی بیشتری در آن قابل رویت است. روش bicubic به دلیل ذات درونیاب، کمترین کیفیت را در نتایج داشته و عدم بکل‌گیری آموزش تقویتی احتمال ایجاد عوارض مصنوعی نادرست را در نتایج افزایش داده است.

در ادامه به منظور ارزیابی بصری میزان حفظ محتوای مکانی در نتایج تولید شده، تصاویر مربوط به بزرگی گرادیان برای روش پیشنهادی و روش‌های مقایسه‌ای در قالب برش‌هایی در دو شکل (۴-۱۰ و ۴-۱۱) برای نتایج کسب شده از تصاویر سنتینل-۲ و لندست-۸ ارائه شده است.



شکل ۴-۱۰: مقایسه بصری تصاویر گرادیان بمنظور ارزیابی بصری میزان حفظ محتوای مکانی در روش‌های پیشنهادی و مقایسه‌ای مربوط به سنجنده سنتینل-۲؛ دو سطر بالا مجموعه داده شهر گرگان؛ دو سطر پایین مجموعه داده شهر یزد



شکل ۴-۱۱: مقایسه بصری تصاویر گرادیان بمنظور ارزیابی بصری میزان حفظ محتوای مکانی در روش‌های پیشنهادی و مقایسه‌ای مربوط به سنجنده لندست-۸؛ دو سطر بالا مجموعه داده شهر گرگان؛ دو سطر پایین مجموعه داده شهر یزد

مقایسه بصری تصاویر بزرگی گرادیان نیز حاکی از حفظ بیشتر محتوای لبه‌های تصویری در روش پیشنهادی نسبت به روش‌های مقایسه‌ای است. در دو شکل (۴-۱۰) و (۴-۱۱) ستون اول (سمت چپ) تصویر پانکروماتیک سنجنده کارتوست (P5) بعنوان یک تصویر با توان تفکیک مکانی بالا و هم‌رده با وضوح تصاویر تولید شده ارائه شده است. مقایسه نتایج بدست آمده از روش پیشنهادی و روش‌های مقایسه‌ای با این تصویر می‌تواند شاهدهی بر کفایت هر



روش در میزان حفظ محتوای مکانی باشد. مجدداً بکارگیری روش bicubic بدترین نتایج را در بازیابی محتوای لبه‌های تصویری داشته و عدم بکارگیری آموزش تقویتی در شبکه Real-ESRGAN کماکان توانایی بازیابی لبه‌های بلوک متراک را نداشته است.

در ادامه ارزیابی نتایج، کمیت‌های مرتبط با سنجش میزان حفظ محتوای طیفی و مکانی (در بخش ۳-۷ معرفی شده‌اند) بمنظور مقایسه کمی نتایج تولید و در جدول (۴-۶) گزارش شده‌اند.

جدول ۶-۴: مقایسه کمی نتایج روش پیشنهادی و روش‌های مقایسه‌ای از منظر میزان حفظ محتوای طیفی و مکانی

معیارهای مکانی			معیارهای طیفی			روش	سنجنده	منطقه
Corr (1)	Compl (1)	SCC (1)	CC (1)	SSIM (1)	PSNR ( $\infty$ )			
۰/۵۳	۰/۶۸	۰/۶۶	۰/۹۹	۰/۴۴۹	۱۸/۵۹	پیشنهادی	سنتینل-۲	یزد
۰/۴۲۹	۰/۴۵۵	۰/۴۴۱	۰/۹۹	۰/۵۰	۱۸/۴۳۶	Bicubic		
۰/۴۴۴	۰/۴۵۹	۰/۴۵۳	۰/۴۹۲	۰/۴۳۵	۱۷/۴۷۶	Real-ESRGAN		
۰/۵۰	۰/۶۴	۰/۵۸	۰/۹۹	۰/۶۱	۱۹/۴۸۴	پیشنهادی	لندست-۸	
۰/۴۳۹	۰/۴۴۸	۰/۴۳۴	۰/۹۹	۰/۴۶۰	۱۹/۹۴	Bicubic		
۰/۴۴۲	۰/۴۵۱	۰/۴۴۱	۰/۴۹۷	۰/۴۴۶	۱۹/۴۱۱	Real-ESRGAN		
۰/۵۱	۰/۶۶	۰/۶۱	۰/۹۹	۰/۴۴۲	۱۵/۴۴۹	پیشنهادی	سنتینل-۲	گرگان
۰/۴۳۷	۰/۴۴۳	۰/۴۴۷	۰/۹۹	۰/۴۴۴	۱۵/۵۳	Bicubic		
۰/۴۴۰	۰/۴۵۶	۰/۴۳۹	۰/۴۹۶	۰/۴۴۱	۱۴/۴۹۶	Real-ESRGAN		
۰/۴۹	۰/۶۰	۰/۵۷	۰/۹۹	۰/۵۶	۱۹/۵۳	پیشنهادی	لندست-۸	
۰/۴۳۲	۰/۴۴۴	۰/۴۴۲	۰/۹۹	۰/۴۵۵	۱۹/۴۴۹	Bicubic		
۰/۴۳۹	۰/۴۴۷	۰/۴۳۵	۰/۴۹۷	۰/۴۵۰	۱۸/۴۹۸	Real-ESRGAN		

مطابق با نتایج مندرج در جدول (۴-۶)، روش پیشنهادی و دوروش مقایسه‌ای از منظر میزان حفظ محتوای طیفی تمایز محسوسی ندارد. از آنجا که روند تلفیق محتوای طیفی به تصاویر با وضوح بهتر در تمامی روش‌های مورد استفاده مشابه است؛ نمی‌توان انتظار تمایزات فاحشی را در نتایج کسب شده از این منظر داشت. روش bicubic به دلیل ماهیت درونیاب خود احتمال تولید محتوای رادیومتریکی غیرواقعی را در تولید تصویر با وضوح بیشتر ندارد و

به همین دلیل بطور متوسط کمترین میزان اعوجاجات طیفی را کسب نموده است. با اینحال، روش پیشنهادی نیز با رقابتی بسیار نزدیک در نیمی از کمیت‌های مرتبط با سنجش میزان حفظ محتوای طیفی گوی سبقت را از روش bicubic ربوده است. عدم بکل‌گیری آموزش تقویتی به کمک تصاویر گوگل ارث منجر به کسب نتایج ضعیف‌تری در تمامی نتایج کسب شده از معیارهای طیفی شده است. این موضوع به دلیل احتمال بیشتر در تولید محتوای رادیومتریکی کاذب قبل از آموزش تقویتی شبکه عصبی GAN می‌باشد. از منظر میزان حفظ محتوای مکانی، رویکرد پیشنهادی در تمامی مجموعه‌های داده و در تمامی شاخص‌های مرتبط با سنجش میزان حفظ محتوای مکانی وضعیت بهتری را نسبت به روش‌های مقایسه‌ای کسب کند. این موضوع علاوه بر نتایج کمی بصورت کیفی نیز در اشکال ارائه شده در این بخش مشهود بود. بطور متوسط در مورد شاخص SCC، مقادیر ثبت شده برای روش پیشنهادی ۳۲/۸۵ درصد از بهترین نتایج کسب شده در روش‌های مقایسه‌ای بالاتر بوده است. این موضوع در حالی است که عدم بکل‌گیری یادگیری تقویتی ظرفیت کاهش کیفیت مکانی را بطور متوسط تا ۳۰/۹ درصد برخوردار است. از این رو به نظر می‌رسد بکل‌گیری رویکرد پیشنهادی بمنظور آموزش تقویتی یک سازوکار موثر در ارتقاء کارایی شبکه‌های عصبی GAN در تولید تصاویر با وضوح مکانی بالاتر باشد.

نتایج تجربی کسب شده از اجرای روش پیشنهادی حاکی از آن است که بکل‌گیری رویکردهای مبتنی بر یادگیری عمیق و شبکه‌های عصبی قادرند تصاویر با وضوح مکانی بهتر را از نسخه‌های با وضوح کمتر تولید سازند. با اینحال کیفیت نتایج بطور مستقیم به کیفیت آموزش شبکه وابسته است. استفاده از شبکه‌های عصبی آموزش یافته به کمک مناطق جغرافیایی دیگر احتمال تولید محتوای تصویری کاذب را افزایش داده و نمی‌توان از شبکه‌های آموزش یافته (به شکل تقویتی) در یک منطقه، انتظار تولید تصاویر مطلوب را در منطقه دیگر داشت. به طور کلی در تصاویر با وضوح بهتر سطح جزئیات مکانی کمتری در مقایسه با تصاویر واقعی با وضوح مشابه مشاهده می‌شود و کاربران چنین روش‌هایی نباید انتظار تولید محتوای مکانی غیرقابل رویت در نسخه با وضوح پایین‌تر را در نسخه‌های با وضوح مکانی بالاتر داشته باشند. مزیت استفاده از روش پیشنهادی را می‌توان در بارزسازی لبه‌های تصویری در

تحلیل‌های مکانی عنوان کرد. این موضوع می‌تواند در کاربردهای کارتوگرافی و تولید نقشه‌های موضوعی با وضوح بالاتر مورد استفاده قرار گیرد.

#### **۴-۴- جمع بندی**

در این فصل به پیاده سازی روش پیشنهادی بر روی چهار مجموعه داده از دو شهر گرگان و یزد پرداخته شد. در این راستا آزمون‌های متفاوتی برای ارزیابی روش پیشنهادی ارائه شد و عملکرد شبکه در شرایط متفاوت به صورت بصری و عددی بررسی گشت. مقایسه بصری در قالب شکل‌های (۴-۶ تا ۴-۱۱) مورد بررسی قرار گرفت و ارزیابی عددی به کمک شش معیار ارزیابی کیفیت معرفی شده در بخش (۳-۷) در جدول (۴-۶) ارائه و تحلیل شد.

## فصل ۵: جمع بندی

### ۵-۱- بحث و نتیجه گیری

در این پژوهش راهکاری دو مرحله ای به منظور بهبود وضوح تصاویر چندطیفی از طریق تولید تصویر پانکروماتیک به کمک تکنیک سوپررزولوشن و سپس تلفیق تصویر پانکروماتیک حاصل شده با باند های تصویر چندطیفی پیشنهاد شد. در راستای تولید تصاویر با وضوح مکانی بهتر از یک شبکه عصبی تخصصی از پیش آموزش یافته استفاده شد و یادگیری تقویتی آن به کمک تصاویر محلی مستخرج از سامانه گوگل-ارث انجام صورت پذیرفت. محدودیت های فناوری به دلیل فقدان داده های آموزشی مورد نظر به دلیل استفاده از سامانه گوگل-ارث در تولید محتوا برای آموزش تقویتی جبران شده است. فقدان محتوای طیفی قابل مقایسه با تصاویر چندطیفی امکان بهبود همزمان وضوح مکانی و طیفی را در شبکه های عصبی فراهم نیاورده و از این رو در گام دوم از روش پیشنهادی، تصاویر تولید شده با وضوح مکانی بهتر با تصاویر چندطیفی تلفیق شده اند. این موضوع ضمن حفظ محتوای طیفی، منجر به تولید تصاویری با وضوح مکانی بهتر از تصاویر چندطیفی اولیه شده است. توسعه راهکاری به منظور بازیابی محتوای طیفی همزمان با بهبود وضوح مکانی در یک سامانه هوش مصنوعی و روش های یادگیری عمیق افق تحقیقاتی پیش رو در این حوزه محسوب می شود. نتایج تجربی کسب شده از اجرای روش پیشنهادی حاکی از آن است که بکارگیری رویکردهای مبتنی بر یادگیری عمیق و شبکه های عصبی قادرند تصاویر با وضوح مکانی بهتر را از نسخه های با وضوح کمتر تولید سازند. با اینحال کیفیت نتایج بطور مستقیم به کیفیت آموزش شبکه وابسته

است. استفاده از شبکه‌های عصبی آموزش یافته به کمک مناطق جغرافیایی دیگر احتمال تولید محتوای تصویری کاذب را افزایش داده و نمی‌توان از شبکه‌های آموزش یافته (به شکل تقویتی) در یک منطقه، انتظار تولید تصاویر مطلوب را در منطقه دیگر داشت. به طور کلی در تصاویر با وضوح بهتر سطح جزئیات مکانی کمتری در مقایسه با تصاویر واقعی با وضوح مشابه مشاهده می‌شود و کاربران چنین روش‌هایی نباید انتظار تولید محتوای مکانی غیر قابل رویت در نسخه با وضوح پایینتر را در نسخه‌های با وضوح مکانی بالاتر داشته باشند. مزیت استفاده از روش پیشنهادی را می‌توان در بارزسازی لبه‌های تصویری در تحلیل‌های مکانی عنوان کرد. این موضوع می‌تواند در کاربردهای کارتوگرافی و تولید نقشه‌های موضوعی با وضوح بالاتر مورد استفاده قرار گیرد.

## ۵-۲- پیشنهادها

به کمک روش پیشنهاد شده در این تحقیق می‌توان وضوح تصاویر اخذ شده از سنجنده‌هایی همچون لندست-۸ و سنتینل-۲ را بهبود داد که این امر در بسیاری از کاربردها می‌تواند کارآمد باشد. لازمه انجام کارهایی همچون شناسایی ساختمان و یا شناسایی تغییرات ساختمانی و شهری، وجود تصاویری دو زمانه با وضوح مطلوب می‌باشد. دستیابی به چنین تصاویری یک چالش محسوب می‌شود که روش پیشنهادی می‌تواند یکی از پاسخ‌های احتمالی به چنین چالش‌هایی باشد. تصاویر سنجش از دوری مانند سنتینل-۲ یا لندست-۸ به راحتی و بدون نیاز به پرداخت هزینه به صورت رایگان در دسترس عموم قرار دارند که افزایش وضوح مکانی و کاربردی کردن چنین تصاویری می‌تواند باعث کاهش چشمگیر هزینه در راستای شناسایی تغییرات ساختمانی یا شناسایی ساختمان باشد.

از طرفی داده‌های سنجش از دوری سنجنده سنتینل-۲ دارای دوره اخذ تصویر در حدود ۵ روز هستند که این امر می‌تواند در زمان وقوع حادثه چنین تصاویری را بسیار ارزشمند نماید. هنگامی که یک بلای طبیعی رخ می‌دهد وجود جفت تصویر قبل و بعد از حادثه با وضوح مطلوب می‌تواند در ارزیابی اولیه خسارات وارد شده و شناسایی سریع مناطق تخریب شده بسیار کارآمد باشد.

در این تحقیق سعی شد با استفاده از الگوریتم های هوش مصنوعی موجود به نتایج قابل قبولی دست یافت که با وجود پیشرفت بسیار سریع و چشم گیر این الگوریتم ها، میتوان چشم انداز بسیار مطلوبی را در این مسیر متصور بود. به طوری که تصاویر پانکروماتیک با دقت بالاتری استخراج شده و به طبع آن نتیجه تلفیق با تصویر چند طیفی دقیق تر بوده و حاوی جزئیات بیشتری خواهد بود.

### ۵-۳- قدر دانی و تشکر

اینجانب در راستای پیشبرد اهداف پژوهش از سازمان جغرافیایی و نیروهای مسلح به دلیل در اختیار قرار دادن تصاویر ماهواره ای کارتوست (P5) کمال تشکر خود را اعلان می دارند.

- [۱] غ. گیلانده, عطا, پورنصرت, عزیزی, ایمانی, "ارزیابی پویایی فضای سبز شهری با استفاده از تکنیک سنجش از راه دور و سیستم اطلاعات جغرافیایی (مطالعه موردی: شهر اردبیل)", فصلنامه علوم و تکنولوژی محیط زیست, Fvol. 23, no. 3, pp. 253-266, 2021
- [۵] ع. دوست, فاخره, مباشری, آبکار, "روشی کارا برای افزایش توأمان قدرت تفکیک طیفی و مکانی در تصاویر ماهواره ای در محیط شهری", نشریه سنجش از دور و GIS ایران, Fvol. 4, no. 4, 2013
- [۲۰] ش. حسینی, سیدی, ح. الهی, رضوان, "ارزیابی مقایسه ای روش های تلفیق داده های سنجش از دوری فراطیفی و پانکروماتیک در مناطق شهری و کشاورزی", نشریه علمی مهندسی نقشه برداری و اطلاعات مکانی, Fvol. 10, no. 2, pp. 63-78, 2019
- [۲۷] بشیرپور, ولدان زوج, محمدجواد, مقصودی, "مقایسه روش های تلفیق مبتنی بر PCA و IHS به منظور تلفیق تصاویر هایپریون و Cartosat-1", نشریه سنجش از دور و GIS ایران, Fvol. 8, no. 4, pp. 17-30, 2016
- [۳۴] م. ن. نیازی, م. زاده, and س. زاده, "ارائه روش ادغام تصاویر چند طیفی و پانکروماتیک IHS-GA مبتنی بر مناطق بهبودیافته گیاهی", نشریه علمی علوم و فنون نقشه برداری, Fvol. 6, no. 1, pp. 235-248, 2016

[۳۶] ک. زاده, رنگزن, کاظم, محمدی, شاهین, "کاربرد تلفیق تصاویر ماهواره‌ای لندست-۸ و سنتینل-۲ در

پایش محیطی, "سنجش‌ازدور و سامانه اطلاعات جغرافیایی در منابع طبیعی, Fvol. 9, no. 3, pp. 53-71, 2018,

[۳۸] ی. ترکی, ص. نژاد, جعفری, "روشی جهت تعیین پارامتر بهینه‌ی فلیترهای نرم به‌منظور تلفیق تصاویر در

فضای فرکانس, "علوم و فناوری فضایی, Fvol. 14, no. 3, pp. 23-37, 2021,

[۳۹] کلانتری, صادق, ع. فرد, احمدی, صادق, "افزایش تفکیک‌پذیری تصویر با استفاده از مدل لبه‌ی تحلیلی,"

نشریه مهندسی برق و الکترونیک ایران, Fvol. 15, no. 2, pp. 45-54, 2018,

[۶۶] بشیرپور, و. زوج, مقصودی, "تلفیق تصاویر سنجش از دوری با روش FFT-PCA," نشریه علمی-پژوهشی

مهندسی فناوری اطلاعات مکانی, Fvol. 5, no. 2, pp. 123-140, 2017,

[2] B. Ye, S. Tian, J. Ge, and Y. Sun, "Assessment of WorldView-3 data for lithological mapping," Remote Sensing, vol. 9, no. 11, p. 1132, 2017.

[3] D. P. Roy, J. Li, H. K. Zhang, and L. Yan, "Best practices for the reprojection and resampling of Sentinel-2 Multi Spectral Instrument Level 1C data," Remote sensing letters, vol. 7, no. 11, pp. 1023-1032, 2016.

[4] M. Gargiulo, A. Mazza, R. Gaetano, G. Ruello, and G. Scarpa, "Fast super-resolution of 20 m Sentinel-2 bands using convolutional neural networks," Remote Sensing, vol. 11, no. 22, p. 2635, 2019.

[6] K. Jia, X. Wang, and X. Tang, "Image transformation based on learning dictionaries across image spaces," IEEE transactions on pattern analysis and machine intelligence, vol. 35, no. 2, pp. 367-380, 2012.

[7] C. Ledig et al., "Photo-realistic single image super-resolution using a generative adversarial network," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 4681-4690.

[8] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part IV 13, 2014: Springer, pp. 184-199.

[9] K. Zhang, J. Liang, L. Van Gool, and R. Timofte, "Designing a practical degradation model for deep blind image super-resolution," in Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 4791-4800.

[10] Y. Huang, S. Li, L. Wang, and T. Tan, "Unfolding the alternating optimization for blind super resolution," Advances in Neural Information Processing Systems, vol. 33, pp. 5632-5643, 2020.



- [11] 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.
- [12] G. Vivone, R. Restaino, M. Dalla Mura, G. Licciardi, and J. Chanussot, "Contrast and error-based fusion schemes for multispectral image pansharpening," IEEE Geoscience and Remote Sensing Letters, vol. 11, no. 5, pp. 930-934, 2013.
- [13] D. Liu and J. Li, "Data field modeling and spectral-spatial feature fusion for hyperspectral data classification," Sensors, vol. 16, no. 12, p. 2146, 2016.
- [14] B. Guo, "Entropy-mediated decision fusion for remotely sensed image classification," Remote Sensing, vol. 11, no. 3, p. 352, 2019.
- [15] N. Yokoya, C. Grohnfeldt, and J. Chanussot, "Hyperspectral and multispectral data fusion: A comparative review of the recent literature," IEEE Geoscience and Remote Sensing Magazine, vol. 5, no. 2, pp. 29-56, 2017.
- [16] B. Aiazzi, S. Baronti, and M. Selva, "Improving component substitution pansharpening through multivariate regression of MS  $\pm$  Pan data," IEEE Transactions on Geoscience and Remote Sensing, vol. 45, no. 10, pp. 3230-3239, 2007.
- [17] J. Liu, "Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details," International Journal of Remote Sensing, vol. 21, no. 18, pp. 3461-3472, 2000.
- [18] J. Choi, H. Park, and D. Seo, "Pansharpening using guided filtering to improve the spatial clarity of VHR satellite imagery," Remote Sensing, vol. 11, no. 6, p. 633, 2019.
- [19] B. Lin, X. Tao, M. Xu, L. Dong, and J. Lu, "Bayesian hyperspectral and multispectral image fusions via double matrix factorization," IEEE Transactions on Geoscience and Remote Sensing, vol. 55, no. 10, pp. 5666-5678, 2017.
- [21] L. Loncan et al., "Hyperspectral pansharpening: A review," IEEE Geoscience and remote sensing magazine, vol. 3, no. 3, pp. 27-46, 2015.
- [22] P. Chavez, S. C. Sides, and J. A. Anderson, "Comparison of three different methods to merge multiresolution and multispectral data- Landsat TM and SPOT panchromatic," Photogrammetric Engineering and remote sensing, vol. 57, no. 3, pp. 295-303, 1991.
- [23] T. Maurer, "How to pan-sharpen images using the gram-schmidt pan-sharpen method—A recipe," The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 40, pp. 239-244, 2013.
- [24] S. Han, H. Li, and H. Gu, "The study on image fusion for high spatial resolution remote sensing images," Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XXXVII. Part B, vol. 7, pp. 1159-1164, 2008.
- [25] K. Rokni, A. Ahmad, K. Solaimani, and S. Hazini, "A new approach for surface water change detection: Integration of pixel level image fusion and image classification techniques," International Journal of Applied Earth Observation and Geoinformation, vol. 34, pp. 226-234, 2015.
- [26] S. Klonus and M. Ehlers, "Performance of evaluation methods in image fusion," in 2009 12th International Conference on Information Fusion, 2009: IEEE, pp. 1409-1416.
- [28] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," in Readings in computer vision: Elsevier, 1987, pp. 671-679.

- [29] Y. Ling, M. Ehlers, E. L. Usery, and M. Madden, "FFT-enhanced IHS transform method for fusing high-resolution satellite images," *ISPRS Journal of photogrammetry and Remote Sensing*, vol. 61, no. 6, pp. 381-392, 2007.
- [30] Y. Zhang and G. Hong, "An IHS and wavelet integrated approach to improve pan-sharpening visual quality of natural colour IKONOS and QuickBird images," *Information fusion*, vol. 6, no. 3, pp. 225-234, 2005.
- [31] B. Chen, B. Huang, and B. Xu, "Multi-source remotely sensed data fusion for improving land cover classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 124, pp. 27-39, 2017.
- [32] Q. Wei, J. Bioucas-Dias, N. Dobigeon, and J.-Y. Tourneret, "Hyperspectral and multispectral image fusion based on a sparse representation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 7, pp. 3658-3668, 2015.
- [33] B. A. Johnson, R. Tateishi, and N. T. Hoan, "Satellite image pansharpening using a hybrid approach for object-based image analysis," *ISPRS International Journal of Geo-Information*, vol. 1, no. 3, pp. 228-241, 2012.
- [35] G. Kaplan and U. Avdan, "Sentinel-2 pan sharpening—comparative analysis," in *Proceedings*, 2018, vol. 2, no. 7: MDPI, p. 345.
- [37] J. Raimundo, S. Lopez-Cuervo Medina, J. F. Prieto, and J. Aguirre de Mata, "Super resolution infrared thermal imaging using pansharpening algorithms: Quantitative assessment and application to UAV thermal imaging," *Sensors*, vol. 21, no. 4, p. 1265, 2021.
- [40] Z. Wei and K.-K. Ma, "Contrast-guided image interpolation," *IEEE Transactions on image processing*, vol. 22, no. 11, pp. 4271-4285, 2013.
- [41] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE transactions on image processing*, vol. 10, no. 10, pp. 1521-1527, 2001.
- [42] J. Schennings, "Deep convolutional neural networks for real-time single frame monocular depth estimation," ed, 2017.
- [43] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*, 2009: IEEE, pp. 248-255.
- [44] X. Ren, L. Bo, and D. Fox, "Rgb-(d) scene labeling: Features and algorithms," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2012: IEEE, pp. 2759-2766.
- [45] R. Socher, Y. Bengio, and C. D. Manning, "Deep learning for NLP (without magic)," in *Tutorial Abstracts of ACL 2012*, 2012, pp. 5-5.
- [46] A. Garcia-Garcia, S. Orts-Escolano, S. Oprea, V. Villena-Martinez, and J. Garcia-Rodriguez, "A review on deep learning techniques applied to semantic segmentation," *arXiv preprint arXiv:1704.06857*, 2017.
- [47] L. Deng, G. Hinton, and B. Kingsbury, "New types of deep neural network learning for speech recognition and related applications: An overview," in *2013 IEEE international conference on acoustics, speech and signal processing*, 2013: IEEE, pp. 8599-8603.
- [48] X. X. Zhu et al., "Deep learning in remote sensing: A comprehensive review and list of resources," *IEEE geoscience and remote sensing magazine*, vol. 5, no. 4, pp. 8-36, 2017.
- [49] M. Lin, Q. Chen, and S. Yan, "Network in network," *arXiv preprint arXiv:1312.4400*, 2013.
- [50] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, "Deep learning for visual understanding: A review," *Neurocomputing*, vol. 187, pp. 27-48, 2016.

- [51] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," arXiv preprint arXiv:1511.07122, 2015.
- [52] Y. Liu, B. Fan, L. Wang, J. Bai, S. Xiang, and C. Pan, "Semantic labeling in very high resolution images via a self-cascaded convolutional neural network," *ISPRS journal of photogrammetry and remote sensing*, vol. 145, pp. 78-95, 2018.
- [53] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, 2012.
- [54] V. Dumoulin and F. Visin, "A guide to convolution arithmetic for deep learning," arXiv preprint arXiv:1603.07285, 2016.
- [55] S. Wager, S. Wang, and P. S. Liang, "Dropout training as adaptive regularization," *Advances in neural information processing systems*, vol. 26, 2013.
- [56] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *International conference on machine learning*, 2015: pmlr, pp. 448-456.
- [57] A. G. Howard, "Some improvements on deep convolutional neural network based image classification," arXiv preprint arXiv:1312.5402, 2013.
- [58] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *IEEE Computer graphics and Applications*, vol. 22, no. 2, pp. 56-65, 2002.
- [59] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in *2009 IEEE 12th international conference on computer vision*, 2009: IEEE, pp. 349-356.
- [60] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 2, pp. 295-307, 2015.
- [61] W. Xu, X. Guangluan, Y. Wang, X. Sun, D. Lin, and W. Yirong, "High quality remote sensing image super-resolution using deep memory connected network," in *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, 2018: IEEE, pp. 8889-8892.
- [62] C. Lanaras, J. Bioucas-Dias, S. Galliani, E. Baltsavias, and K. Schindler, "Super-resolution of Sentinel-2 images: Learning a globally applicable deep neural network," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 146, pp. 305-319, 2018.
- [63] 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, no. 13, p. 1588, 2019.
- [64] M. Habibi, A. Ahmadyfard, and H. Hassanpour, "Single Image Super-Resolution via Learning Segmented Regions of the Input Image," *Journal of Machine Vision and Image Processing*, vol. 7, no. 1, pp. 111-121, 2020.
- [65] Y. Li and B. Li, "Super-Resolution of Sentinel-2 Images at 10m Resolution without Reference Images," 2021.
- [67] I. Goodfellow et al., "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, 2014.
- [68] 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.
- [69] Y. Zhu and S. Newsam, "Densenet for dense flow," in *2017 IEEE international conference on image processing (ICIP)*, 2017: IEEE, pp. 790-794.

- [70] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 4700-4708.
- [71] X. Wang, L. Xie, C. Dong, and Y. Shan, "Real-esrgan: Training real-world blind super-resolution with pure synthetic data," in Proceedings of the IEEE/CVF international conference on computer vision, 2021, pp. 1905-1914.
- [72] X. Wang et al., "Esrgan: Enhanced super-resolution generative adversarial networks," in Proceedings of the European conference on computer vision (ECCV) workshops, 2018, pp. 0-0.
- [73] S. Bharati, P. Podder, M. Mondal, and V. Prasath, "CO-ResNet: Optimized ResNet model for COVID-19 diagnosis from X-ray images," International Journal of Hybrid Intelligent Systems, vol. 17, no. 1-2, pp. 71-85, 2021.
- [74] E. H. Helmer and B. Ruefenacht, "Cloud-free satellite image mosaics with regression trees and histogram matching," Photogrammetric Engineering & Remote Sensing, vol. 71, no. 9, pp. 1079-1089, 2005.
- [75] C. A. Laben and B. V. Brower, "Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening," ed: Google Patents, 2000.
- [76] B. Goodarzi, J. Javidan, and M. J. Dehghani, "New changes of local binary patterns and classification and segmentation of seabed images," Journal of Information and Communication Technology, vol. 27, no. 27, p. 1, 2019.
- [77] S. Talebi, A. Zarea, S. Sadeghian, and H. Arefi, "A Hierarchical Unsupervised Method for Tree Detection Using Aerial Imagery and LiDAR," Geospatial Engineering Journal, vol. 5, no. 3, pp. 55-66, 2014.
- [78] I. Amro, J. Mateos, M. Vega, R. Molina, and A. K. Katsaggelos, "A survey of classical methods and new trends in pansharpening of multispectral images," EURASIP Journal on Advances in Signal Processing, vol. 2011, no. 1, pp. 1-22, 2011.
- [79] Z. Xiang, L. Xiao, J. Yang, W. Liao, and W. Philips, "Detail-Injection-Model-Inspired Deep Fusion Network for Pansharpening," IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-15, 2022.
- [80] C. Keeratikasikorn and I. Trisirisatayawong, "Reconstruction of 30m dem from 90 m SRTM DEM with bicubic polynomial interpolation method," The International Archives of the Photogrammetry, Remote sensing and spatial information Sciences, vol. 37, pp. 791-794, 2008.

## **Abstract**

Improving the spatial resolution of multispectral images is one of the common pre-processing steps in reaching the maximum potential of these images in remote sensing applications. The presence of images with higher spatial resolution along with multispectral images allows the process of improving spatial resolution to be performed through image pan-sharpening methods. The lack of simultaneous panchromatic image sensors in the satellite platforms imposes challenges related to co-registration and asynchronies when using images of different satellite sensors in the process of image pan-sharpening. In such a situation, super-resolution techniques are considered as alternative approaches to improve spatial resolution. Using the generative adversarial network (GAN) is one of the effective methods in this field that require the existence of multiple training data. Generally, it is not possible to prepare two satellite images with the same spectral resolution and different spatial resolution from a specific region that is required for training the network. Therefore, in this research, an approach with two main steps is designed to improve the spatial resolution of multispectral images. In the first step, a deep super-resolution generative adversarial network is used to improve the resolution of the true color composition of multispectral images. A boosting strategy is exploited to deeply train the GAN network using the resampled images extracted from the Google-Earth. In the second step, the spectral contents are added to the super-resolution images through the traditional pan-sharpening method. The results demonstrated that the proposed approach improved the spatial resolution of multispectral images by 32.85% better than the best comparative method and maintained the spectral content without the need to provide extensive training data.

## **Keywords**

Image fusion, Super-resolution, Multispectral images, Deep learning, Boost learning.

In the name of God



Ministry of Science, Research, and Technology  
Tafresh University

**Civil Engineering Department**

**MSc student of photogrammetry, Department of Geodesy and Surveying Engineering**

Title

# **Spatial enhancement of medium resolution satellite imagery through deep learning methods**

By

**Hamid Amini Amirkolae**

Supervisor

**Dr. Dr. Alireza Safdarinezhad**

Advisor

**Dr. Dr. Hamed Amini Amirkolae**

End Month & Year: october 2023

