نشریه علمی مهندسی فناوری اطلاعات مکانی

سال یازدهم، شماره دوم، تابستان ۱۴۰۲

Vol.11, No.2, Summer 2023

۴۱ -۶۲

مقاله پژوهشی

DOR: 20.1001.1.20089635.1402.11.2.3.6

[Downloaded from jgit.kntu.ac.ir on 2026-01-01]

[DOR: 20.1001.1.20089635.1402.11.2.3.6]

[DOI: 10.61186/jgit.11.2.41]

بهبود وضوح مکانی تصاویر چندطیفی فضایی در نبود تصاویر با توان تفکیک مکانی بالاتر به کمک روش‌های تلفیق تصویر و یادگیری عمیق

حمید امینی امیرکلائی[۱]، علیرضا صفدری نژاد[۲*]، حامد امینی امیرکلائی[۳]

۱- دانشجوی کارشناسی ارشد فتوگرامتری، گروه ژئودزی و مهندسی نقشه‌برداری، دانشگاه تفرش

۲- استادیار گروه ژئودزی و مهندسی نقشه‌برداری، دانشگاه تفرش

۳- دکتری مهندسی نقشه‌برداری گرایش فتوگرامتری، دانشکده مهندسی نقشه برداری و اطلاعات مکانی، دانشگاه تهران

تاریخ دریافت مقاله: ۱۴۰۲/۰۳/۲۰ تاریخ پذیرش مقاله: ۱۴۰۲/۰۵/۲۸

چکیده

بهبود وضوح مکانی تصاویر چندطیفی یکی از پیش‌پردازش‌های رایج در استفاده حداکثری از ظرفیت این تصاویر در کاربردهای سنجش‌ازدوری است. وجود تصاویر با وضوح مکانی بهتر در کنار تصاویر چندطیفی امکان انجام این فرایند را از طریق روش‌های تلفیق تصاویر فراهم می‌آورند. عدم اخذ همزمان تصاویر با وضوح مکانی بالاتر در یک سکوی فضایی، چالش‌های مربوط به هم‌مرجع‌سازی و عدم هم‌زمانی را هنگام بکارگیری تصاویر سایر سنجنده‌های فضایی به فرایند تلفیق تصاویر تحمیل می‌سازد. در چنین شرایطی، تولید تصاویر به کمک روش‌های سوپررزولوشن از رویکردهای جایگزین برای افزایش وضوح مکانی محسوب می‌شوند. استفاده از شبکه‌های عصبی کانولوشنی یکی از روش‌های موثر 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 این رو، در این تحقیق روند بهبود وضوح مکانی تصاویر چندطیفی از طریق روشی دو مرحله‌ای به اجرا رسیده است. در گام نخست از یک شبکه عصبی کانولوشنی برای تولید نسخه واضح‌تر رنگی از باندهای مرئی تصاویر چندطیفی استفاده شده که آموزش تقویتی آن به کمک تصاویر بازآرایی شده از بستر گوگل-ارث تامین می‌گردد. در گام دوم، تصویر با وضوح مکانی بهتر در نقش یک تصویر پانکروماتیک با تصویر چندطیفی تلفیق شده و محتوای طیفی به آن افزوده می‌شود. نتایج نشان داد که رویکرد پیشنهادی بدون نیاز به تامین پرهزینه داده‌های آموزشی ضمن تامین شرایط برابر در حفظ محتوای طیفی، توان بهبود وضوح مکانی تصاویر چندطیفی را تا ۳۲/۸۵ درصد بهتر از بهترین روش مقایسه‌ای دارد.

کلید واژه‌ها : تلفیق تصاویر، سوپررزولوشن، تصاویر چند طیفی، یادگیری عمیق، آموزش تقویتی.

* نویسنده مکاتبه کننده: استان مرکزی، شهرستان تفرش، کیلومتر ۵ جاده تهران، دانشگاه تفرش.

تلفن: ۰۹۱۲-۳۲۵۸۸۵۲

Email: safdarinezhad@tafreshu.ac.ir

۱- مقدمه

امروزه سامانه‌های سنجش‌ازدوری متعددی به‌منظور پایش زمین با توان تفکیک مکانی، طیفی و زمانی مختلف توسعه‌یافته‌اند. از دیدگاه فنی، توان تفکیک مکانی و طیفی سنجنده‌های فضایی ارتباط معکوسی دارند. علیرغم گستردگی کاربردهای این تصاویر، وضوح مکانی متوسط و یا کم آن‌ها همواره به‌عنوان یک عامل محدودکننده در برخی از کاربردها به شمار می‌رود. روش‌های محاسباتی مختلفی به‌منظور بهبود توان تفکیک مکانی تصاویر چندطیفی توسعه‌یافته‌اند که عموماً به یک تصویر با توان تفکیک مکانی بالاتر نیازمندند. در برخی از سکوهای فضایی امکان اخذ هم‌زمان تصاویر پانکروماتیک با توان تفکیک مکانی بالاتر در کنار تصاویر چندطیفی وجود داشته و این موضوع امکان تلفیق آن‌ها را به سهولت فراهم می‌آورد. بااین‌حال، برخی از سنجنده‌های 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. محدودیت‌هایی از قبیل عدم هم‌زمانی تصاویر 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[۱] نام‌گذاری می‌شوند، بر پایه بیشینه‌سازی هم‌زمان مزایای طیفی تصاویر چندطیفی و مکانی تصاویر پانکروماتیک در تولید محصول نهایی بناشده است. محصول این روش‌ها، تصاویر با توان تفکیک مکانی مشابه با تصویر پانکروماتیک و محتوای طیفی شبیه به تصاویر چندطیفی است. تاکنون این رویکرد به شکل‌های مختلفی در حوزه تلفیق تصاویر مورداستفاده قرارگرفته است.

[۱] *Pan-Sharpening*

علیدوست و همکاران (۲۰۱۳) با استفاده از اطلاعات مکانی تصویر آیکونوس[۲] با توان تفکیک مکانی ۴ متر و اطلاعات طیفی تصویر هایپریون[۳] 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، مقایسه‌ای بین پیکسل بازسازی‌شده تصویر هایپریون و پیکسل واقعی صورت دادند [۱]. محمد نژاد نیازی و همکاران (۲۰۱۶) روشی جدید بجای روش مشهور شدت-رنگ-اشباع (*IHS*[۴]) باهدف کاهش اعوجاجات مکانی در تصویر شدت[۵] برای مناطق گیاهی پیشنهاد نمودند. در این روش از شاخص گیاهی تفاضلی نرمال شده جهت شناسایی مناطق پوشش گیاهی استفاده شده است. آنالیزهای بصری و آماری حاکی از آن است که روش مورد بحث توانسته کیفیت تصویر تلفیق‌شده را نسبت به خروجی روش‌های مرسوم بهبود دهد [۲]. بشیر پور و همکاران (۲۰۱۷) مقایسه‌ای بین روش‌های تلفیق تصاویر ۳۰ متری هایپریون و تصاویر پانکروماتیک ۲/۵ متری کارتوست-۱[۶] انجام دادند. این مقایسه بین روش‌های تبدیل سریع فوریه (*FFT*[۷]) مبتنی بر *IHS* [۳] ، موجک[۸] مبتنی بر *IHS* [۴] و تحلیل مؤلفه‌های اصلی (*PCA*[۹]) مبتنی بر موجک [۵] صورت گرفته است. نتایج عددی نشان داد روش‌های مبتنی بر *PCA* دقت بالاتری در تلفیق تصاویر نسبت به *IHS* داشته‌اند [۵]. شاه‌حسینی و همکاران (۲۰۱۹) به

[۲] *IKONOS*
[۳] *Hyperion*
[۴] *Intensity Hue Saturation*
[۵] *Intensity*
[۶] *Cartosat-1*
[۷] *Fast Fourier Transform*
[۸] *Wavelet*
[۹] *Principal Component Analysis*

[Downloaded from jgit.kntu.ac.ir on 2026-01-01]
[DOR: 20.1001.1.20089635.1402.11.2.3.6]
[DOI: 10.61186/jgit.11.2.41]

تحلیل مقایسه‌ای دقت و کارایی روش‌های نوین تلفیق داده در مناطق کشاورزی و شهری با استفاده از تصاویر فراطیفی با انواع کلاس‌های پیچیده و متنوع به‌صورت بصری و کمی پرداختند. ایشان در روند تلفیق از تصویر فراطیفی هایپریون و تصویر پانکروماتیک مربوط به سنجنده ALI[۱] استفاده نمودند. نتایج نشان داد که روش مبتنی بر تجزیه ماتریس به روش فاکتورگیری دوگانه غیرمنفی ($CNMF$[۲]) عملکرد بهتری در مقایسه با سایر روش‌ها دارد، به‌طوری‌که توانسته توان تفکیک عوارض را در سطح پیکسل با جزئیات حفظ کند اما نیازمند پردازش زیادی است. همچنین روش غیرخطی شدت-رنگ-اشباع[۳] زمان پردازش کمتری داشته اما در حفظ جزئیات عملکرد مناسبی نشان نداده است [۶]. کابلی-زاده و همکاران (۲۰۱۹) به پایش شهر اهواز، از تلفیق تصاویر سنجش‌ازدور لندست-۸[۴] و سنتینل-۲ به کمک روش‌های تلفیق تصاویر، گرام اشمیت[۵]، نرمال‌سازی رنگ (CN[۶])، برووی[۷]، تبدیل موجک گسسته (DWT[۸])، برهم‌نهی شدت مبتنی بر فیلترهای نرم ($SFIM$[۹]) و الگوریتم ضرب[۱۰] پرداختند. نتایج نشان داد که روش‌های DWT و برووی به ترتیب بهترین روش ازلحاظ حفظ اطلاعات طیفی و مکانی بوده و بیشترین اعوجاجات ازنظر طیفی مربوط به الگوریتم ضرب می‌باشد. ایشان نشان دادند که روش $SFIM$ از جنبه طیفی و مکانی دارای دقت مناسب است [۷]. یعقوبی و همکاران (۲۰۲۱) از فیلترگذاری نرم با ابعاد بهینه در فضای فرکانس به‌منظور بهبود تلفیق تصاویر چندطیفی و پانکروماتیک استفاده نمودند. در این راستا از شاخصی وزن‌دار متشکل از دو جزء رنگ‌پذیری و جزئیات مکانی استفاده نمودند که توانسته با اتکا به معیارهای هندسی و آماری فضای ویژگی در کنار فیلترهای نرم، نتایج مطلوبی را در شناسایی ابعاد بهینه فیلترگذاری داشته باشد [۸].

در روش‌های تلفیق تصویر، بهبود توان تفکیک مکانی نیازمند دسترسی هم‌زمان به تصویر با توان تفکیک مکانی بالا (تصویر پانکروماتیک) و تصویر چندطیفی متناظر بوده که تأمین آن همیشه ممکن نیست. روش‌هایی بانام سوپررزولوشن[۱۱] یا بهبود 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 سوپررزولوشن اولین بار در سال ۱۹۸۰ مورد مطالعه قرار گرفت و پس از آن روش‌های گوناگونی در چهار دهه گذشته توسعه‌یافته است. در ادامه برخی از روش‌های توسعه‌یافته در این حوزه مرور شده است [۹]. یکی از قدیمی‌ترین روش‌ها در این حوزه، روش‌های مبتنی بر درون‌یابی است. در این روش‌ها عموماً از یک عملگر پایه درون‌یابی برای بازسازی نقاط خالی شبکه منظم مربوط به تصویر با وضوح بالاتر، استفاده می‌شود. مهم‌ترین مزیت این روش‌ها سادگی و پیچیدگی محاسبات کم است. روش درون‌یابی نزدیک‌ترین همسایه، ساده‌ترین و سریع‌ترین روش در میان این روش‌ها قلمداد شده که باعث تخریب همواری تصویر و رخداد حالت شطرنجی در نتایج می‌گردد [۱۰]. برای حل این مسئله روش‌های درون‌یابی با مرتبه نرمی بالاتر پیشنهاد شده (مانند درون‌یابی‌دوخطی و مکعبی) که برای تخمین درجات خاکستری در تصاویر ارتقاء

[۱] *Advance Land Imager*

[۲] *Coupled Non-negative Matrix Factorization*

[۳] *Non-linear Intensity Hue-Saturation*

[۴] *LandSat-8*

[۵] *GramSchmidt*

[۶] *Color Normalized*

[۷] *Brovey*

[۸] *Discrete Wavelet Transform*

[۹] *Smoothing Filter-based Intensity Modulation*

[۱۰] *Multiplication*

[۱۱] *Super Resolution*

[Downloaded from jgit.kntu.ac.ir on 2026-01-01]
[DOR: 20.1001.1.20089635.1402.11.2.3.6]
[DOI: 10.61186/jgit.11.2.41]

یافته از تعداد بیشتری همسایه از تصویر مبنا استفاده می‌نمایند. این موضوع کاهش اثر شطرنجی شدن را به دنبال داشته اما بروز ماتی در محل لبه‌ها را به دنبال دارد [۱۱]. برای کاهش این چالش نیز روش‌های درون‌یابی جهتی توسعه‌یافته که با مشخص ساختن راستای مناسب در هر نقطه، روند نمونه‌برداری را به اجرا می‌رسانند. در این روش‌ها راستای درون‌یابی موازی با لبه‌های تصویری انتخاب‌شده که این موضوع حفظ محتوای لبه‌ها را به دنبال خواهد شد [۱۲]. دسته‌ای دیگر، روش‌های مبتنی بر یادگیری می‌باشند. در این روش‌ها با در اختیار داشتن تعداد بسیار زیادی از تصاویر آموزشی متناظر در دو نسخه با وضوح بالا (*HR*[۱]) و وضوح پایین (*LR*[۲])، تلاش می‌شود رابطه‌ای بین این تصاویر از طریق روش‌های یادگیری برقرار گردد. مدل‌های آموزش‌یافته در ادامه برای افزایش وضوح تصویر *LR* استفاده می‌شوند [۱۲]. برخی از روش‌های پیشرفته به‌کاررفته در این حوزه عبارت‌اند از: شبکه مولد متخاصم (*GAN*[۳]) [۱۳]، شبکه عصبی کانولوشنی سوپررزولوشن (*SRCNN*[۴]) [۱۴]، شبکه مولد تخاصمی سوپر رزولوشن نابینا (*BSRGAN*[۵]) [۱۵]، شبکه عمیق کانولوشنی بازگشتی (*DRCN*[۶]) [۱۶]، شبکه عمیق کانولوشنی در سطح زیر پیکسل (*ESPCN*[۷]) [۱۷] و شبکه سوپررزولوشن خودآموز (*SESR*[۸]) [۱۸]. در ادامه برخی از تحقیقات مرتبط با به‌کارگیری روش‌های سوپررزولوشن مرور شده است.

[۱] *High Resolution*

[۲] *Low Resolution*

[۳] *Generative Adversarial Network*

[۴] *Super-Resolution Convolutional Neural Network*

[۵] *Blind image Super-Resolution*

[۶] *Deeply-Recursive Convolutional Network*

[۷] *Efficient Sub-Pixel Convolutional Neural network*

[۸] *Self-Exemplars Super-Resolution*

فریمن[۹] و همکاران (۲۰۰۲) یک شبکه مارکوفی 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 استفاده‌شده که این موضوع منجر به حفظ جزئیات ظریف تصویری گردیده است [۱۹]. گلسنر[۱۰] و همکاران همکاران (۲۰۰۹) نیز یک روش 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 تولیدشده، سپس به ازای پنجره‌هایی با ابعاد معلوم (مثلاً ۷×۷) از تصویر اولیه، موقعیت‌های شبیه در تصویر با مقیاس پایین‌تر جست‌وجو می‌شود. در گام بعدی، موقعیت‌های متناظر پنجره‌های شناسایی‌شده در تصویر با مقیاس بزرگ‌تر، به‌عنوان داده متناظر در روند آموزش یک شبکه افزایش وضوح مکانی انتخاب می‌شوند [۲۰]. دانگ[۱۱] و همکاران (۲۰۱۵) نیز یک روش یادگیری عمیق مبتنی بر شبکه عصبی کانولوشنی (*CNN*[۱۲]) برای بهبود وضوح مکانی به کمک تنها یک تصویر، پیشنهاد دادند. آن‌ها یک نگاشت انتها به انتها بین تصویر اولیه و تصویر *LR* تولیدشده از آن، در غالب 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 به‌صورت غیرخطی به نمایش‌هایی با وضوح بالاتر تبدیل

[۹] *Freeman*

[۱۰] *Glasner*

[۱۱] *Dong*

[۱۲] *Convolutional Neural Network*

[Downloaded from jgit.kntu.ac.ir on 2026-01-01]
[DOR: 20.1001.1.20089635.1402.11.2.3.6]
[DOI: 10.61186/jgit.11.2.41]

می‌شوند. آخرین لایه مسئول ادغام نقشه‌های لایه‌های قبلی با هدف تولید یک تصویر با وضوح بالاتر است. 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 [۲۱]. هوک[۱] و همکاران (۲۰۱۹)، تکنیک‌های سوپررزولوشن تصویر مبتنی بر یادگیری عمیق با استفاده از شبکه عصبی کانولوشن (*CNN*) و شبکه متخاصم مولد (*GAN*) را برای افزایش چهار برابری وضوح تصاویر سنجش‌ازدور پیشنهاد دادند. نتایج تجربی نشان داد که مدل *GAN* عملکرد بهتری نسبت به مدل‌های *CNN* داشته است [۲۲]. حبیبی و همکاران (۲۰۲۰) 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 یادگیری می‌شود. برای ایجاد تخمین بهتری از تصویر خروجی، دو تصویر *HR* و *LR* آموزشی را بر اساس خوشه‌بندی ناحیه‌بندی نموده و مدل‌های رگرسیونی را برای هر ناحیه به‌طور مجزا آموزش داده‌اند. از طرفی برای کاهش اثر تاری در لبه‌های تصویر خروجی، مدل‌های جداگانه‌ای برای یادگیری لبه‌ها ارائه‌شده است [۲۳]. لی و همکاران (۲۰۲۱) از شبکه عصبی با هدف سوپررزولوشن تصویر برای تبدیل باندهای (۲۰ متری و ۶۰ متری تصاویر سنتینل-۲) به ۱۰ متری استفاده نمودند. آن‌ها از یک شبکه آموزش‌دیده کانولوشنال پیشرفته برای ارتقا توان تفکیک مکانی باندهای با وضوح ۲۰ و ۶۰ متری استفاده نمودند. داده‌های مورداستفاده به‌منظور آموزش در سطح گسترده‌ای از مکان‌های جغرافیایی انتخاب‌شده تا شبکه آموزش‌یافته از قابلیت تعمیم‌پذیری مناسبی در مناطق مختلف آب‌وهوایی و در انواع پوشش زمین برخوردار باشد. این شبکه ویژگی‌های طیفی را به‌خوبی حفظ می‌کند و نتایج بصری قانع‌کننده‌ای در وضوح ۱۰ متری ارائه می‌دهد [۲۴]. عدم دسترسی به تصاویر *HR* در سنجنده‌های فاقد باند پانکروماتیک مشوقی برای به‌کارگیری روش‌های سوپررزولوشن در تولید تصاویر چندطیفی با وضوح بالاتر است. استفاده از روش‌های یادگیری عمیق در این حوزه، ایده اصلی این تحقیق تلقی می‌شود. با این حال دسترسی به تصاویر مناسب برای آموزش چنین شبکه‌هایی یک چالش اجرایی محسوب می‌شود. به عبارت بهتر، تأمین هم‌زمان تصاویر چندطیفی 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 مقاله راهکاری دو مرحله‌ای به‌منظور افزایش وضوح مکانی تصاویر چندطیفی در زمان فقدان باند طیفی پانکروماتیک ارائه شده است. این روند ضمن افزایش توان تفکیک مکانی تصاویر اولیه، محتوای طیفی را مشابه با روش‌های سنتی تلفیق تصاویر بازیابی می‌نماید. توسعه راهکاری باهدف تولید تصاویر چندطیفی با وضوح مکانی بالاتر به‌طور مستقیم توسط شبکه‌های عصبی عمیق نیازمند دسترسی به تصاویر چندطیفی در دو نسخه با وضوح مکانی متفاوت جهت استفاده در روند آموزش است. عدم دسترسی به تصاویر چندطیفی با وضوح مکانی بالا و یا هزینه‌های مترتب برای دسترسی به چنین منابع داده یکی از انگیزه‌های توسعه راهکار دو مرحله‌ای در روش پیشنهادی بوده است. عدم نیاز به تصاویر چندطیفی با توان تفکیک مکانی بالاتر، استفاده از ظرفیت تصاویر ماهواره‌ای منتشرشده در بستر گوگل-ارث در روند یادگیری عمیق و بازیابی محتوای طیفی بر پایه روش‌های سنتی تلفیق از ویژگی‌های این روش به شمار می‌روند.

۲-روش پیشنهادی

امروزه شبکه‌های عصبی آموزش‌یافته به کمک تکنیک‌های یادگیری عمیق در حوزه‌های مختلفی از

[۱] *Hoque*

[Downloaded from jgit.kntu.ac.ir on 2026-01-01]
[DOR: 20.1001.1.20089635.1402.11.2.3.6]
[DOI: 10.61186/jgit.11.2.41]

پردازش تصاویر سنجش‌ازدوری استفاده می‌شوند. مطابق با منابع مرور شده در این مقاله، روش‌های متنوعی به‌منظور ارتقاء وضوح مکانی تصاویر سنجش‌ازدوری به کمک تکنیک‌های یادگیری عمیق توسعه‌یافته است. در تمامی رویکردهای مبتنی بر یادگیری عمیق دسترسی به یک پایگاه داده غنی از نمونه‌های آموزشی پیش‌شرط موفقیت این روش‌ها بیان‌شده است [۱۳]. از منظر کاربردی، دستیابی به چنین محتوای اطلاعاتی با چالش‌های اجرایی روبروست. به‌عنوان مثال در سنجنده ماهواره‌ای سنتینل-۲ که فاقد باند پانکروماتیک است؛ تأمین داده‌های آموزشی با محتوای طیفی مشابه و توان تفکیک مکانی بالاتر توأم با مشکلات و هزینه‌های اجرایی است. به‌کارگیری تصاویر ماهواره‌ای *HR* به‌عنوان تصاویر متناظر، عدم تطابق در باندهای طیفی و چالش‌های هم‌مرجع‌سازی را به دنبال دارند. به عبارت بهتر، عدم مشابهت موقعیت، عرض و تابع پاسخ طیفی سنجنده‌های چندطیفی امکان تأمین داده‌های آموزشی چندطیفی در دو نسخه *HR* و *LR* را با محدودیت مواجه می‌سازد. از این رو در این مقاله ایده‌ای کم‌هزینه برای تولید تصاویر *HR* چندطیفی پیشنهاد شده است.

گوگل-ارث (*Earth- Google*) 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 به‌ندرت از آن بی‌بهره بوده‌اند. تاکنون راهکارهای متنوعی به‌منظور ارتقاء بهره‌وری از تصاویر این پایگاه داده توسعه‌یافته است. ایجاد نرم‌افزارهای دانلود و زمین‌مرجع‌سازی دقیق تصاویر گوگل مانند نرم افزار *StitchMapper* و یا *SASPlanet* نمونه نرم‌افزارهایی تلقی شده که باهدف دانلود و بهبود دقت مکانی تصاویر گوگل مبتنی بر روش‌های پردازش‌تصویری توسعه‌یافته‌اند. ایده این تحقیق، آموزش عمیق شبکه‌های عصبی بر پایه به‌کارگیری 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 عمیق یک شبکه عصبی کانولوشنی استفاده می‌شوند. این تصاویر صرفاً در محدوده مکانی تصاویر چندطیفی و با کمترین اختلاف‌زمانی 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 که نقش تصویر پانکروماتیک را در ادامه ایفا می‌نماید. در گام دوم، تصاویر با وضوح مکانی بالاتر به‌دست‌آمده از نتایج یادگیری عمیق با باندهای طیفی تصاویر چندطیفی از طریق روش‌های مرسوم تلفیق می‌شوند. محصول نهایی در این روند بدون نیاز به تصاویر با وضوح مکانی و طیفی بالا و صرفاً با بهره‌گیری از تصاویر منتشرشده در بستر گوگل-ارث تولید می‌شوند. روندنمای ارائه‌شده در شکل (۱) روند کلی روش پیشنهادی را نشان می‌دهد.

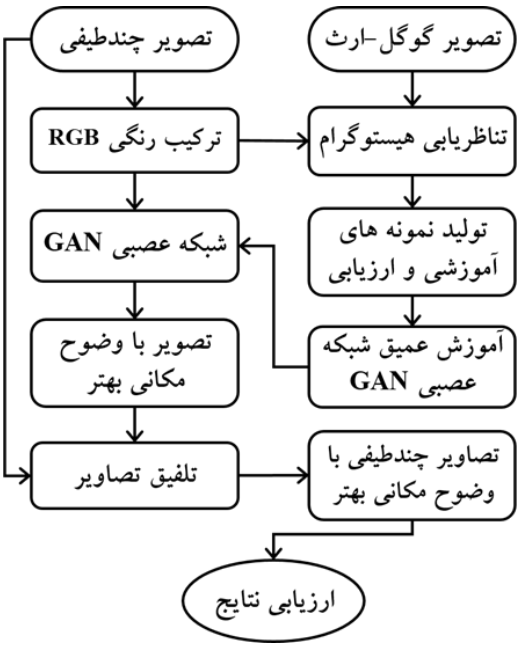
مطابق با شکل (۱)، به ازای هر تصویر چندطیفی، تصاویر *HR* متناظر از بستر گوگل-ارث استخراج شده و در دو نسخه متناسب با توان تفکیک مکانی تصویر چندطیفی و تصویر ارتقاءیافته مکانی مورد انتظار، نمونه‌برداری کاهشی[۱] می‌گردد. قبل از این اقدام، شباهت رادیومتریکی تصاویر گوگل به تصاویر چندطیفی از طریق روش تناظریابی هیستوگرام[۲] افزایش می‌یابد. پس از تنظیم وضوح مکانی و تناظریابی رادیومتریکی تصاویر مستخرج از سامانه گوگل-ارث، نمونه‌های آموزشی و ارزیابی به‌منظور آموزش عمیق یک شبکه عصبی کانولوشنی تخاصمی آماده‌سازی می‌شود. این شبکه قادر خواهد بود که توان تفکیک

[۱] *Down sample*

[۲] *Histogram matching*

[Downloaded from jgit.kntu.ac.ir on 2026-01-01]
[DOR: 20.1001.1.20089635.1402.11.2.3.6]
[DOI: 10.61186/jgit.11.2.41]

مکانی را تا چهار برابر ارتقاء بخشد. بازیابی محتوای طیفی در این روند به دلیل فقدان مابه‌ازای طیفی مناسب در روند آموزش میسر نبوده و این روند صرفاً باهدف تولید محتوای مشابه با تصاویر پانکروماتیک با وضوح مکانی بالا به اجرا می‌رسد. ازاین‌رو، شبکه آموزش‌یافته به کمک ترکیب رنگی واقعی از تصاویر چندطیفی تغذیه شده و تصاویر با وضوح مکانی بالاتر را تولید می‌سازد. این تصاویر پس از تبدیل به نسخه تک باندی، با تصاویر چندطیفی اولیه تلفیق‌شده و محصول ارتقاءیافته را تولید می‌کنند. در ادامه این بخش هر یک از اجزای این روند در قالب زیر بخش‌های مجزا به تفصیل تشریح شده‌اند.



شکل ۱: روندنمای کلی روش پیشنهادی

۲-۱- شبکه عمیق *GAN*

مطابق با محتوای مندرج در روندنمای روش پیشنهادی شکل(۱) تولید تصاویر با توان تفکیک مکانی بالا به‌منظور تلفیق با تصاویر چندطیفی به اجرا رسیده است. برای این منظور از شبکه عمیق سوپررزولوشن تخاصمی بهبودیافته (*Real-ESRGAN*[۱]) استفاده‌شده است که یک الگوریتم سوپررزولوشن مبتنی بر یادگیری عمیق است و می‌تواند وضوح مکانی تصاویر با وضوح مکانی پایین را افزایش دهد [۲۵]. *Real-ESRGAN* یک نسخه پیشرفته از *ESRGAN* [۲۶] بوده که یک الگوریتم قدرتمند سوپررزولوشن بر اساس معماری *GAN* [۱۳] است. *Real-ESRGAN* برای افزایش وضوح مکانی چهار برابری تصویر از یک فرآیند دو مرحله‌ای استفاده می‌کند. به‌طوری‌که ابتدا با استفاده از شبکه مولد[۲] یک تصویر اولیه با وضوح مکانی بالا تولید می‌کند. سپس از شبکه تفکیک‌کننده[۳] برای مقایسه تصویر به‌دست‌آمده با تصویر مرجع دارای وضوح مکانی بالا به‌منظور شناسایی تفاوت‌ها استفاده می‌کند. در ادامه، از این تفاوت‌ها برای به‌روزرسانی شبکه مولد

[۱] *Enhanced Super-Resolution Generative Adversarial Networks*

[۲] *Generator*

[۳] *Discriminator*

[Downloaded from jgit.kntu.ac.ir on 2026-01-01]
[DOR: 20.1001.1.20089635.1402.11.2.3.6]
[DOI: 10.61186/jgit.11.2.41]

استفاده شده تا تفاوت بین تصویر تولیدشده و تصویر مرجع را به حداقل برساند.

شبکه مولد دارای دو بخش کدگذار و کدگشا است. بخش کدگذار شبکه مولد از تکنیکی به نام یادگیری باقیمانده استفاده می‌کند که به شبکه اجازه می‌دهد به‌جای 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 می‌تواند عملکرد کلی شبکه را بهبود بخشد و به کاهش میزان داده‌های آموزشی موردنیاز کمک کند [۲۵]. شبکه مولد از بلوک‌های شبکه پیچشی متراکم (*DensNet*) [۲۷] تشکیل‌شده است که این بلوک‌ها مطابق با معماری به‌کار رفته در شبکه عصبی باقیمانده (*ResNet*) [۲۸] به هم متصل شده‌اند. به این صورت که یک اتصال مستقیم بین ورودی بلوک هر *DensNet* و خروجی آن در هر بار اعمال این بلوک ایجادشده است (این قسمت با حرف *a* در شکل (۲) نمایش یافته است). هر بلوک *DensNet* خود از اعمال توالی لایه‌های کانولوشن (*Conv*)، لایه فعال‌ساز ریلو (*ReLU*) و *Conv* در دفعات متفاوت تشکیل‌شده است که در آن خروجی یک‌بار اعمال این توالی به نتایج حاصل از اعمال سایر توالی‌ها متصل است (این قسمت با حرف *b* در شکل(۲) نمایش یافته است).

در بخش کدگشای شبکه مولد با اعمال لایه‌های *Conv*، لایه پخش‌کننده پیکسل[۱] و *ReLU* سعی می‌شود که با بهره‌گیری از ویژگی‌های به‌دست‌آمده از بخش کدگذاری شامل ویژگی‌های سطح پایین (لبه‌ها و بافت‌ها) تا ویژگی‌های سطح بالا مانند عوارض، ویژگی‌هایی مؤثر با وضوح مکانی ۲ برابر تولید شود. در ادامه با تکرار فرایند مطرح‌شده، وضوح مکانی ویژگی‌ها بار دیگر دو برابر شده و نهایتاً با اعمال یک‌لایه *Conv* تعداد کانال‌های ویژگی‌ها کاهش‌یافته و تصویر نهایی با وضوح مکانی ۴ برابر به دست می‌آید (شکل (۲))

در شبکه تفکیک‌کننده با اعمال لایه‌های *Conv*، لایه نرمال‌سازی دسته‌ای (*BN*[۲]) و *ReLU* و بهره‌گیری از رابطه (۱) سعی می‌شود که واقعی یا جعلی بودن تصویر تولیدشده توسط شبکه مولد تعیین گردد.

$$D(I_f, I_f) = \sigma\left(C(I_r) - E\left[C(I_f)\right]\right) \to 1$$
$$D(I_f, I_f) = \sigma\left(C(I_f) - E\left[C(I_r)\right]\right) \to 0 \qquad \text{رابطه (۱)}$$

دررابطه(۱)، I_r تصویر واقعی، I_f تصویر جعلی، σ تابع سیگموئید، E تابع میانگین و C خروجی شبکه تفکیک‌کننده می‌باشد. اگر D به ۱ میل کند، نشان‌دهنده واقعی بودن و اگر به صفر میل کند نشان‌دهنده جعلی بودن تصویر تولیدشده است. یکی از ویژگی‌های کلیدی *Real-ESRGAN* تابع اتلاف ادراکی[۳] آن است که برای بهینه‌سازی شبکه مولد استفاده می‌شود. این تابع اتلاف طراحی‌شده است تا نه‌تنها تفاوت بین تصاویر واقعی تولیدشده و تصاویر حقیقی، بلکه شباهت ادراکی آن‌ها را نیز اندازه‌گیری کند. این کار برای اطمینان حاصل کردن از یکسان بودن کیفیت بصری تصاویر تولیدشده با تصاویر حقیقی است.

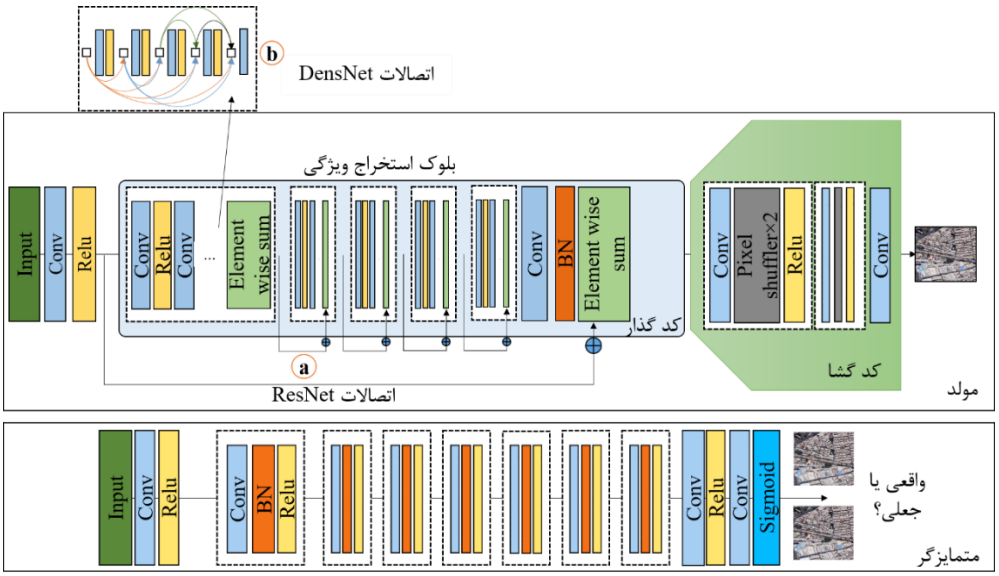
شبکه عمیق مذکور توسط مجموعه داده‌های متعددی آموزش‌دیده و نسخه آموزش‌یافته آن در آدرس (*https://github.com/xinntao/Real-ESRGAN*) در دسترس است. این شبکه عصبی در دو نسخه اولیه (بدون آموزش مجدد) و آموزش‌یافته تقویتی (به کمک داده‌های محدود از منطقه موردنظر) در روند بازسازی تصاویر با وضوح بهتر استفاده شده است.

[۱] *pixel shuffler*

[۲] *Batch Normalization*

[۳] *Perceptual*

[Downloaded from jgit.kntu.ac.ir on 2026-01-01]
[DOR: 20.1001.1.20089635.1402.11.2.3.6]
[DOI: 10.61186/jgit.11.2.41]



شکل ۲: معماری شبکه عمیق *Real-ESRGAN*

۲-۲- آماده‌سازی نمونه‌های آموزشی برای یادگیری عمیق

به‌کارگیری نمونه‌های آموزشی مناسب پیش‌شرط 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 شده‌اند. این موضوع باهدف کاهش هزینه‌های مربوط به تأمین داده‌های آموزشی صورت گرفته است. به عبارت بهتر، داده‌های آموزشی مربوط به آموزش تقویتی شبکه عصبی *GAN* به‌طور کامل با داده‌های موردنیاز در روند به‌کارگیری شبکه عصبی متفاوت است. داده‌های آموزشی از سامانه گوگل استخراج‌شده و در روند آموزش از تصاویر چندطیفی استفاده نمی‌شود. بدین معنی که تصاویر مستخرج از سامانه گوگل-ارث پس از دانلود، در دو نسخه متناسب با وضوح مکانی تصاویر ورودی به شبکه عصبی (*LR*) و تصاویر مورد انتظار پس از افزایش وضوح مکانی (*HR*) مورد نمونه‌برداری مجدد قرار می‌گیرند. به دلیل اینکه هر دو نسخه از تصاویر *LR* و *HR* از یک منبع (تصاویر استخراج‌شده از سامانه گوگل-ارث) این تصاویر به‌صورت ذاتی هم مرجع بوده و ارتباط هندسی میان آن‌ها معلوم است. با این حال، یک پیش‌پردازش رادیومتریکی نیز به‌منظور انطباق بیشتر داده‌های آموزشی با تصاویر چندطیفی 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 چندطیفی از بستر گوگل-ارث، محتوای رادیومتریکی آن‌ها با باندهای رنگی مرئی از تصویر چندطیفی شبیه شدند. این اقدام از طریق اجرای روش تناظریابی هیستوگرام صورت گرفته است. تناظریابی هیستوگرام یک تبدیل نقطه‌ای بوده که شباهت توزیع درجات خاکستری تصویر منبع (گوگل) را به تصویر مرجع (چندطیفی) افزایش می‌دهد [۲۹]. رابطه (۲) تابع تبدیل در فرایند تناظریابی هیستوگرام است.

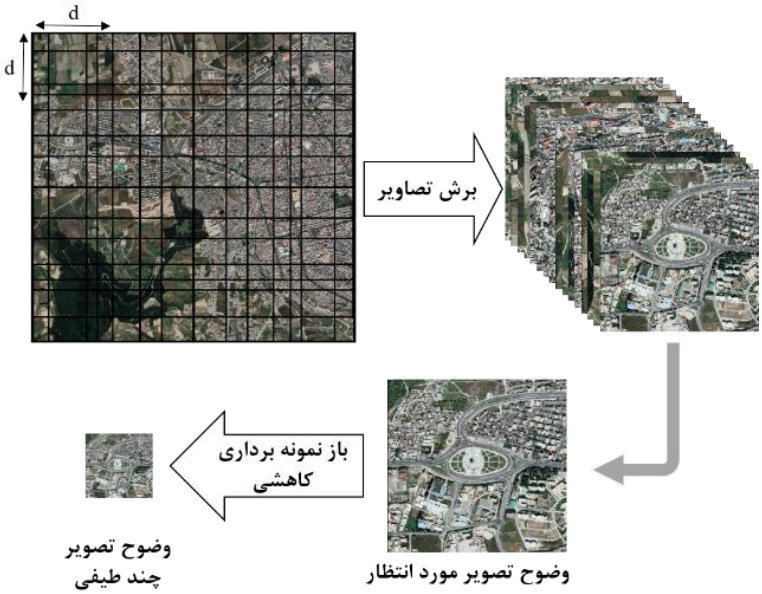
رابطه (۲)

$$y = f_r^{-1}\left(f_s(x)\right)$$

در رابطه (۲)، 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

[Downloaded from jgit.kntu.ac.ir on 2026-01-01]
[DOR: 20.1001.1.20089635.1402.11.2.3.6]
[DOI: 10.61186/jgit.11.2.41]

منبع و y درجه خاکستری اصلاح‌شده پس از تناظریابی هیستوگرام هستند. به‌کارگیری روند تناظریابی هیستوگرام به‌منظور افزایش شباهت رادیومتریکی تصاویر گوگل با تصاویر چندطیفی تنها نقشی بوده که تصاویر چندطیفی در روند آموزش تقویتی شبکه عصبی عمیق ایفا می‌کنند. بدیهی است که در روند تناظریابی هیستوگرام دقت بالا در روند هم‌مرجع‌سازی تصاویر ضروری نبوده و وجود همپوشانی مکانی میان تصاویر مرجع و منبع کفایت لازم برای این روند را تأمین می‌سازد.با توجه به عملکرد شبکه عصبی *Real-ESRGAN* در افزایش چهار برابری وضوح تصاویر و ضرورت تأمین محتوای تصویری مشابه با تصاویر چندطیفی در روند آموزش، وضوح تصاویر مستخرج از گوگل-ارث به ۱- وضوح مکانی تصویر چندطیفی و ۲- وضوح تصویر مورد انتظار پس از بهبود توان تفکیک مکانی، تبدیل می‌شود. شکل (۳) روند شماتیک این اقدام را نشان می‌دهد.



شکل ۳: روند شماتیک باز نمونه‌برداری کاهشی به‌منظور آماده‌سازی داده‌های آموزشی از تصاویر گوگل-ارث

با توجه به روند اتخاذ شده در تولید تصاویر آموزشی شکل (۳)، تصاویر مورداستفاده به‌صورت ذاتی هم مرجع بوده و اقدامی مبنی بر هم‌مرجع‌سازی آن‌ها نیاز نیست. این تصاویر به برش‌های ۱۰۰×۱۰۰ در نسخه هم‌وضوح با تصویر چندطیفی و ۴۰۰×۴۰۰ در نسخه ارتقاءیافته مکانی، برای آموزش عمیق شبکه عصبی معرفی‌شده در بخش (۲-۲) استفاده شده‌اند. در تولید این تصاویر آموزشی، پنجره‌هایی با ابعاد ذکرشده توأم با همپوشانی ۷۵ درصدی از تصاویر گوگل (در نسخه‌های تنظیم‌شده از منظر مکانی) استخراج می‌شود. وجود همپوشانی در نمونه‌های آموزشی ضامن تأمین نمونه‌های آموزشی متنوع و بیشتر در روند آموزش تقویتی خواهد بود.

۲-۳- 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 چندطیفی است. مطابق با روندنمای شکل (۱)، پس از تولید تصاویر با وضوح مکانی بالاتر از تصاویر چندطیفی، این تصاویر با تصاویر چندطیفی تلفیق می‌شوند. در این مرحله از روش گرام

[Downloaded from jgit.kntu.ac.ir on 2026-01-01]
[DOR: 20.1001.1.20089635.1402.11.2.3.6]
[DOI: 10.61186/jgit.11.2.41]

اشمیت[1] به‌عنوان روش تلفیق استفاده‌شده است.
تبدیل گرام اشمیت روشی در جبر خطی برای متعامدسازی مجموعه‌ای از بردارها است. اولین بار کداک از این تبدیل برای تلفیق تصاویر استفاده نمود [۳۰]. در این روش، یک تصویر پانکروماتیک هم‌اندازه با تصویر چندطیفی از طریق میانگین‌گیری وزن‌دار از باندهای طیفی شبیه‌سازی‌شده و سایر باندهای طیفی نسبت به آن و نسبت به یکدیگر از طریق تبدیل گرام-اشمیت متعامدسازی می‌شوند. سپس باند پانکروماتیک اصلی جایگزین باند شبیه‌سازی‌شده 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 به ازای هر تصویر برآورد می‌شود. جزئیات مربوط به اجرای این روش تلفیق توسط مورر (۲۰۱۳) ارائه شده است [۳۱]. علاوه بر نتایج مطلوب گزارش‌شده برای این روش، امکان به‌کارگیری این روش در تصاویر چند و ابرطیفی از مزیت‌های آن قلمداد می‌شود [۳۱].

۲-۴- رویکردهای ارزیابی نتایج

متناسب با سازوکار روش پیشنهادشده در این مقاله، رویکردهای ارزیابی نتایج با تفاوت‌های اندکی شبیه به رویکردهای مرسوم در ارزیابی نتایج تلفیق تصاویر است [۸]. محصول نهایی در روش پیشنهادی یک تصویر چندطیفی 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 قرارگرفته است. تصاویر ماهواره‌ای مورداستفاده 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.
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 چندطیفی از طریق نمونه‌برداری کاهشی به تصویری با توان تفکیک مکانی پایین‌تر تبدیل‌شده و سپس روند بازسازی نسخه با وضوح بالاتر از طریق روش پیشنهادی به اجرا می‌رسد. نتایج تولیدشده تصاویری با توان تفکیک طیفی و مکانی مشابه با تصاویر چندطیفی اولیه بوده که امکان مقایسه میزان اعوجاجات طیفی در آن‌ها وجود خواهد داشت. شکل (۴) این رویکرد را به‌صورت شماتیک نشان می‌دهد.
شاخص‌های متنوعی به‌منظور ارزیابی نتایج تلفیق تصاویر وجود داشته که به دودسته معیارهای طیفی و مکانی قابل تفکیک هستند. در این تحقیق از شاخص‌های ۱- حداکثر نسبت سیگنال به نویز (*PSNR*)[2]، ۲- شباهت ساختاری (*SMI*)[3] و ۳- ضریب همبستگی (*CC*)[4] به‌عنوان شاخص‌های سنجش حفظ اطلاعات طیفی و شاخص‌های ۱- ضریب همبستگی مکانی (*SCC*)، ۲- کامل بودن لبه‌های تصویری و ۳- صحت حفظ محتوای لبه به‌عنوان رویکرد ارزیابی کفایت مکانی نتایج استفاده شده است. دو مورد آخر از شاخص‌های حفظ محتوای مکانی از طریق مقایسه لبه‌های مستخرج بین تصویر تولیدشده و تصویر مبنا در قالب یک ماتریس درهم‌ریختگی دو دویی تولید می‌شوند. جزئیات مرتبط با نحوه محاسبه هر یک از شاخص‌های مذکور در [۸، ۳۲ و ۳۳] منتشر شده است.

[1] *Gram-Schmidt*

[2] *Peak signal-to-noise ratio*

[3] *Structural Similarity Index*

[4] *Cross Correlation*

[Downloaded from jgit.kntu.ac.ir on 2026-01-01]
[DOR: 20.1001.1.20089635.1402.11.2.3.6]
[DOI: 10.61186/jgit.11.2.41]

هم‌مرجع‌سازی تصاویر به‌منظور محاسبه شاخص‌های مکانی ضروری بوده و در این تحقیق، از چندجمله‌ای‌های دوبعدی از درجه دو یا سه (متناسب با میزان اختلافات هندسی) که پارامترهای آن در روند تناظریابی کمترین مربعات برآورد شده‌اند برای هم‌مرجع‌سازی مکانی تصاویر هنگام ارزیابی میزان حفظ محتوای مکانی استفاده شده است.

شکل ۴: روند شماتیک تولید محتوا مورد مقایسه جهت ارزیابی میزان حفظ محتوای طیفی در روش پیشنهادی

۳- داده‌ها و منطقه مورد مطالعه

در این تحقیق دو مجموعه داده مربوط به بخش‌هایی از دو شهر گرگان و یزد جهت پیاده‌سازی روش پیشنهادی استفاده شده است. در دسترس بودن، وجود تنوع در پوشش زمین و بافت شهری متمایز دلایل انتخاب این دو مجموعه داده بوده است. هر مجموعه داده از چهار صحنه تصویر ماهواره‌ای تشکیل شده است. تصویر ماهواره‌ای سنتینل-۲، تصویر ماهواره‌ای لندست-۸، تصویر مستخرج از سامانه گوگل-ارث و تصویر ماهواره‌ای کارتوست-۱ (*P5*)، چهار تصویر هر مجموعه داده را تشکیل می‌دهند. تصاویر گوگل برای آموزش تقویتی شبکه عصبی عمیق *GAN* استفاده‌شده و تصاویر ماهواره‌ای کارتوست *P5* نیز به‌منظور ارزیابی نتایج مربوط به میزان حفظ محتوای مکانی بکار گرفته‌شده‌اند. جزئیات مرتبط با هر یک از تصاویر این دو مجموعه داده در جدول (۱) گزارش‌شده است.

جدول ۱: جزئیات مربوط به مجموعه داده‌های مورد استفاده در ارزیابی نتایج روش پیشنهادی

مجموعه داده	موقعیت جغرافیایی	وسعت (km^2)	تصاویر	تعداد باندهای طیفی	وضوح مکانی (m)
گرگان	$\varphi = ۳۶^{\circ}۴۹'$ $\lambda = ۵۴^{\circ}۲۶'$	۳۲۶	سنتینل -۲	۱۰	۱۰ و ۲۰
			لندست -۸	۹	۱۵ و ۳۰
			کارتوست *P5*	۱	۲/۵
			گوگل ارث	*RGB*	حدود ۱ متر
یزد	$\varphi = ۳۱^{\circ}۵۳'$ $\lambda = ۵۴^{\circ}۲۰'$	۸۸۳	سنتینل -۲	۱۰	۱۰ و ۲۰
			لندست -۸	۹	۱۵ و ۳۰
			کارتوست *P5*	۱	۲/۵
			گوگل ارث	*RGB*	حدود ۱ متر

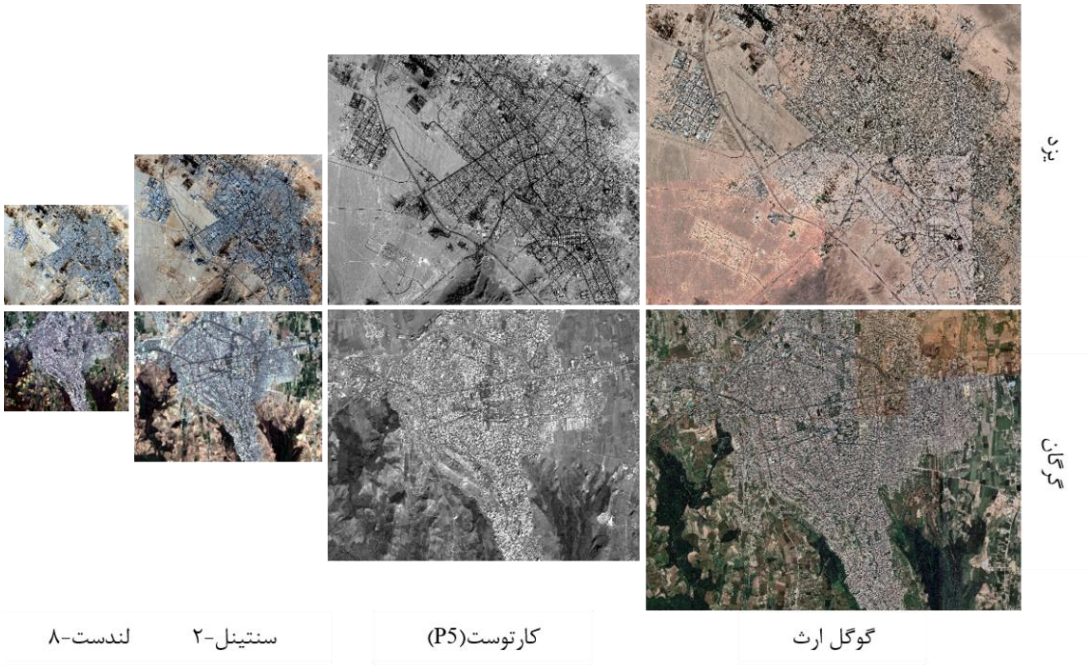
لازم به ذکر است که تصاویر ماهواره‌ای سنتینل-۲ متشکل از ۱۳ باند طیفی با توان تفکیک مکانی ۱۰، ۲۰ و ۶۰ متری بوده که در این تحقیق به دلیل اختلاف زیاد توان تفکیک مکانی باندهای ۶۰ متری نسبت به محصول تولیدی با وضوح مکانی بالا، از آن‌ها استفاده نشده است. اقدام مشابهی در مورد باندهای حرارتی

[Downloaded from jgit.kntu.ac.ir on 2026-01-01]
[DOR: 20.1001.1.20089635.1402.11.2.3.6]
[DOI: 10.61186/jgit.11.2.41]

سنجنده لندست- ۸ نیز به اجرا رسیده است. شکل (۵) بخشی از تصاویر مربوط به هر مجموعه داده را در کنار هم نشان می‌دهد.

۴-پیاده‌سازی و ارزیابی نتایج

در این بخش پیاده‌سازی 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 شده‌اند. به‌عبارت 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 تلفیق‌شده و توان تفکیک مکانی آن‌ها یکسان شده است. به دلیل کیفیت پایین فرایند تلفیق برای باندهای طیفی ۶۰ متری در سنجنده سنتینل-۲ و ۹۰ متری در سنجنده لندست-۸ از بکارگیری این باندهای طیفی صرفنظر شده است. به‌طور موازی، تصاویر مستخرج از سامانه گوگل-ارث از طریق نمونه‌برداری کاهشی به دو مجموعه آموزشی برای تصاویر سنتینل-۲ و لندست-۸ تبدیل شده‌اند. 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.



شکل ۵: تصاویر مورد استفاده از شهر یزد و گرگان

با فرض توان تفکیک مکانی یک متری برای تصاویر مستخرج از سامانه گوگل ارث، تصاویر بازنمونه‌برداری شده برای آموزش تقویتی شبکه عصبی عمیق در مجموعه اول به‌وضوح 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.

در گام سوم از روند پیاده‌سازی، آموزش تقویتی شبکه عصبی *Real-ESRGAN* در دو نسخه به کمک

[Downloaded from jgit.kntu.ac.ir on 2026-01-01]
[DOR: 20.1001.1.20089635.1402.11.2.3.6]
[DOI: 10.61186/jgit.11.2.41]

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 استفاده‌شده است. مطابق با روندنمای ارائه‌شده در شکل (۱)، از شبکه عصبی *Real-ESRGAN* صرفاً باهدف 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 استفاده‌شده و محصول تولیدشده نیز یک تصویر سه باندِ قرمز-سبز-آبی (*RGB*) است؛ محصول تولیدشده از طریق میانگین‌گیری به یک تصویر پانکروماتیک تبدیل می‌گردد. درنهایت، 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*، ۱۶ گیگابایت *RAM* و کارت گرافیک *NVIDIA GeForce GTX 3060 Ti* دارای ۶ گیگابایت حافظه *GPU* به اجرا رسیده است. روش و جزئیات مرتبط با روند آموزش تقویتی شبکه عصبی *Real-ESRGAN* در جدول (۲) گزارش‌شده است [۲۵].

ارزیابی روش پیشنهادی در این تحقیق به کمک دو روش مقایسه‌ای صورت پذیرفته است. عدم به‌کارگیری تکنیک آموزش تقویتی در شبکه *Real-ESRGAN* و به‌کارگیری روش سنتی درون‌یابی دو مکعبی (*Bicubic*) [۳۴] دو راهکار مقایسه‌ای مورداستفاده هستند. شکل (۶) و (۷) به ترتیب نتایج کسب‌شده از روش پیشنهادی برای تصاویر سنتینل-۲ و لندست-۸ را به‌همراه سایر روش‌های مقایسه‌ای برای دو پنجره از مجموعه داده‌های مورداستفاده در این تحقیق نشان می‌دهد. در این تصاویر صرفاً ترکیب‌های رنگی مرئی برای نمایش نتایج استفاده شده‌اند.

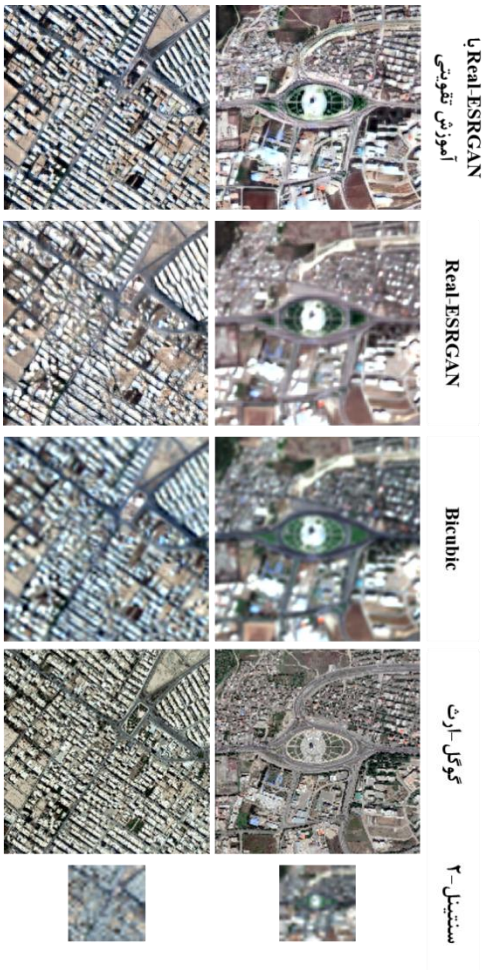
جدول ۲: روش و تنظیمات مرتبط با روند آموزش تقویتی شبکه عصبی *Real-ESRGAN*

روش یادگیری	الگوریتم بهینه‌سازی *ADAM*[۱]
تعداد اپک‌های آموزش	۳۰
نرخ کاهش وزن[۲]	۵×۱۰^{-۳}
نرخ یادگیری	۲×۱۰^{-۴}

[۱] *Adaptive Moment Estimation*

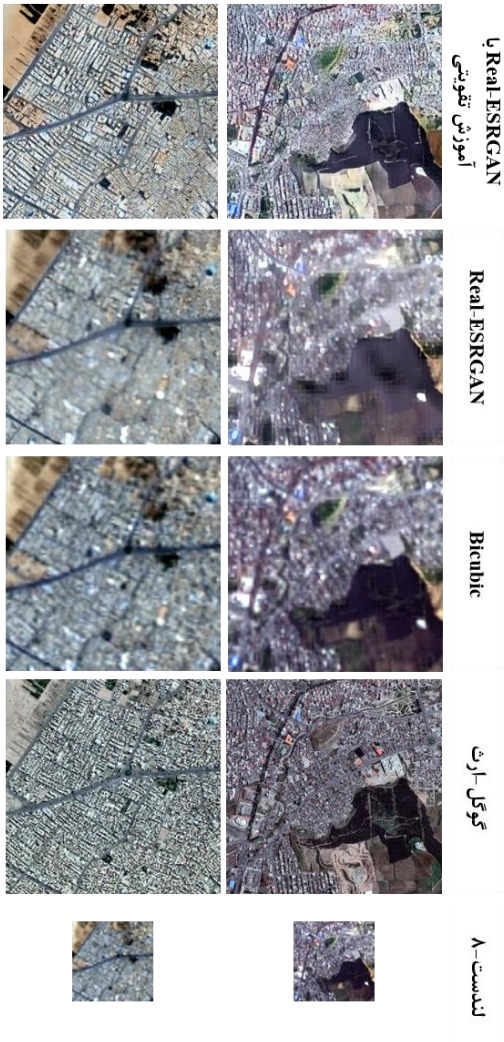
[۲] *Weight decay*

[Downloaded from jgit.kntu.ac.ir on 2026-01-01]
[DOR: 20.1001.1.20089635.1402.11.2.3.6]
[DOI: 10.61186/jgit.11.2.41]



شکل ۶: نمایی از نتایج کسب‌شده از روش پیشنهادی و سایر روش‌های مقایسه‌ای در مورد تصاویر سنجنده سنتینل-۲؛ از راست ستون اول مجموعه داده شهر یزد؛ سطر دوم مجموعه داده شهر گرگان

[Downloaded from jgit.kntu.ac.ir on 2026-01-01]
[DOR: 20.1001.1.20089635.1402.11.2.3.6]
[DOI: 10.61186/jgit.11.2.41]

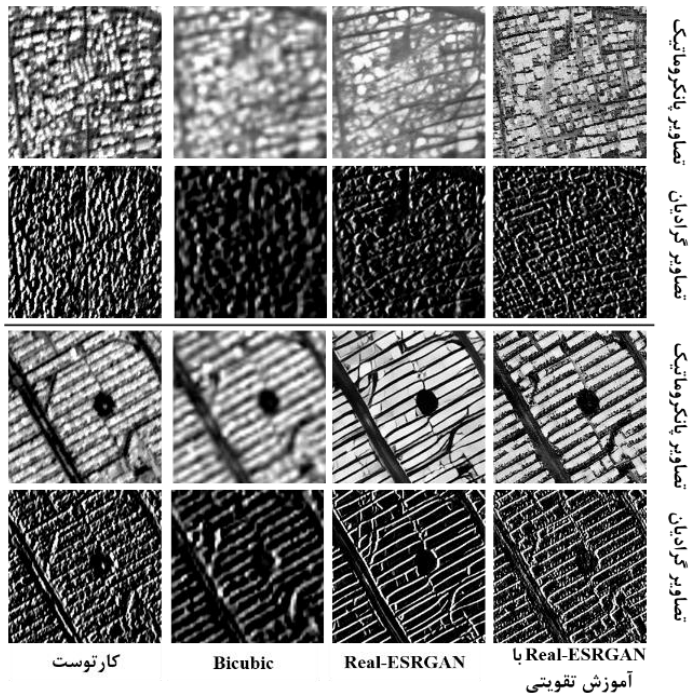


شکل ۷: نمایی از نتایج کسب‌شده از روش پیشنهادی و سایر روش‌های مقایسه‌ای در مورد تصاویر سنجنده لندست-۸؛ از راست ستون اول مجموعه داده شهر یزد؛ سطر دوم مجموعه داده شهر گرگان

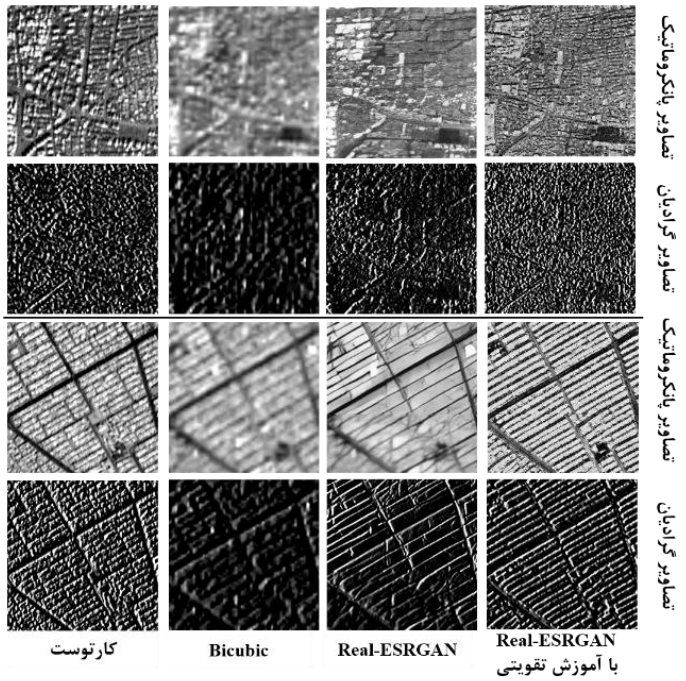
مطابق با نتایج ارائه‌شده در شکل‌های(۶) و (۷)، روش پیشنهادی ازنظر بصری نتایج بهتری را نسبت به دو روش مقایسه‌ای کسب نموده و جزئیات مکانی بیشتری در آن قابل روئیت است. روش درون‌یابی دو مکعبی به دلیل ذات درون‌یاب، کمترین کیفیت را در نتایج داشته و عدم به‌کارگیری آموزش تقویتی احتمال ایجاد عوارض مصنوعی نادرست را در نتایج افزایش داده است.

در ادامه به‌منظور ارزیابی بصری میزان حفظ محتوای مکانی در نتایج تولیدشده، تصاویر مربوط به بزرگی گرادیان برای روش پیشنهادی و روش‌های مقایسه‌ای در قالب برش‌هایی در دو شکل(۸) و (۹) برای نتایج کسب‌شده از تصاویر سنتینل-۲ و لندست-۸ ارائه شده‌اند.

[Downloaded from jgit.kntu.ac.ir on 2026-01-01]
[DOR: 20.1001.1.20089635.1402.11.2.3.6]
[DOI: 10.61186/jgit.11.2.41]



شکل ۸: مقایسه بصری تصاویر گرادیان به‌منظور 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 ۹: مقایسه بصری تصاویر گرادیان به‌منظور 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

[Downloaded from jgit.kntu.ac.ir on 2026-01-01]
[DOR: 20.1001.1.20089635.1402.11.2.3.6]
[DOI: 10.61186/jgit.11.2.41]

مقایسه بصری تصاویر بزرگی گرادیان نیز حاکی از حفظ بیشتر محتوای لبه‌های تصویری در روش پیشنهادی نسبت به روش‌های مقایسه‌ای است.

در دو شکل(۸) و (۹) ستون اول (سمت چپ) تصویر باند پانکروماتیک سنجنده کارتوست (*P5*) به‌عنوان یک تصویر با توان تفکیک مکانی بالا و هم‌رده با وضوح تصاویر تولیدشده، ارائه شده است. مقایسه نتایج به‌دست‌آمده از روش پیشنهادی و روش‌های مقایسه‌ای با این تصویر (کارتوست) می‌تواند شاهدی بر کفایت هر روش در میزان حفظ محتوای مکانی باشد. مجدداً به‌کارگیری روش درون‌یابی دو مکعبی بدترین نتایج را در بازیابی محتوای لبه‌های تصویری داشته و عدم به‌کارگیری آموزش تقویتی در شبکه *Real-ESRGAN* کماکان توانایی بازیابی کامل لبه‌های بلوک ساختمانی متراکم را نداشته است.

در ادامه ارزیابی نتایج، کمیت‌های مرتبط با سنجش میزان حفظ محتوای طیفی و مکانی (در بخش ۲-۴ معرفی‌شده‌اند) به‌منظور مقایسه کمی نتایج تولید و در جدول(۳) گزارش‌شده‌اند.

جدول ۳: مقایسه کمی نتایج روش پیشنهادی و روش‌های مقایسه‌ای از منظر میزان حفظ محتوای طیفی و مکانی

منطقه	سنجنده	روش	معیارهای طیفی			معیارهای مکانی		
			PSNR (∞)	*SSIM* (1)	*CC* (1)	*SCC* (1)	*Compl* (1)	*Corr* (1)
یزد	سنتینل-۲	پیشنهادی	**۱۸٫۵۹**	۰٫۴۹	**۰٫۹۹**	**۰٫۶۶**	**۰٫۶۸**	**۰٫۵۳**
		Bicubic	۱۸٫۳۶	**۰٫۵۰**	**۰٫۹۹**	۰٫۴۱	۰٫۵۵	۰٫۲۹
		Real-ESRGAN	۱۷٫۷۶	۰٫۳۵	۰٫۹۲	۰٫۵۳	۰٫۵۹	۰٫۴۴
	لندست-۸	پیشنهادی	۱۹٫۸۴	**۰٫۶۱**	**۰٫۹۹**	**۰٫۵۸**	**۰٫۶۴**	**۰٫۵۰**
		Bicubic	**۱۹٫۹۴**	۰٫۶۰	**۰٫۹۹**	۰٫۳۴	۰٫۴۸	۰٫۳۹
		Real-ESRGAN	۱۹٫۱۱	۰٫۴۶	۰٫۹۷	۰٫۴۱	۰٫۵۱	۰٫۴۲
گرگان	سنتینل-۲	پیشنهادی	۱۵٫۴۹	۰٫۴۲	**۰٫۹۹**	**۰٫۶۱**	**۰٫۶۶**	**۰٫۵۱**
		Bicubic	**۱۵٫۵۳**	**۰٫۴۴**	**۰٫۹۹**	۰٫۴۷	۰٫۴۳	۰٫۳۷
		Real-ESRGAN	۱۴٫۹۶	۰٫۴۱	۰٫۹۶	۰٫۳۹	۰٫۵۶	۰٫۴۰
	لندست-۸	پیشنهادی	**۱۹٫۵۳**	**۰٫۵۶**	**۰٫۹۹**	**۰٫۵۷**	**۰٫۶۰**	**۰٫۴۹**
		Bicubic	۱۹٫۴۹	۰٫۵۵	**۰٫۹۹**	۰٫۴۲	۰٫۴۴	۰٫۳۲
		Real-ESRGAN	۱۸٫۹۸	۰٫۵۰	۰٫۹۷	۰٫۳۵	۰٫۴۷	۰٫۳۹

مطابق با نتایج مندرج در جدول (۳)، روش پیشنهادی و دو روش مقایسه‌ای از منظر میزان حفظ محتوای طیفی تمایز محسوسی ندارد. از آنجاکه روند تلفیق محتوای طیفی به تصاویر با وضوح بهتر در تمامی روش‌های مورداستفاده مشابه است؛ نمی‌توان انتظار تمایزات فاحشی را در نتایج کسب‌شده از این منظر داشت. روش درون‌یابی دو مکعبی به دلیل ماهیت درون‌یاب خود احتمال تولید محتوای رادیومتریکی غیرواقعی را در تولید تصویر با وضوح بیشتر ندارد و به همین دلیل به‌طور 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 درون‌یابی دو مکعبی

[Downloaded from jgit.kntu.ac.ir on 2026-01-01]
[DOR: 20.1001.1.20089635.1402.11.2.3.6]
[DOI: 10.61186/jgit.11.2.41]

ربوده است. عدم به‌کارگیری آموزش تقویتی به کمک تصاویر گوگل ارث منجر به کسب نتایج ضعیف‌تری در تمامی نتایج کسب‌شده 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 به‌صورت کیفی نیز در اشکال ارائه‌شده در این بخش مشهود بود. به‌طور متوسط در مورد شاخص *SCC*، مقادیر ثبت‌شده برای روش پیشنهادی ۳۲٫۸۵ درصد از بهترین نتایج کسب‌شده در روش‌های مقایسه‌ای بالاتر بوده است. این موضوع در حالی است که عدم به‌کارگیری یادگیری تقویتی ظرفیت کاهش کیفیت مکانی را به‌طور متوسط تا ۳۰٫۹ درصد برخوردار است. از این‌رو به نظر می‌رسد به‌کارگیری رویکرد پیشنهادی به‌منظور آموزش تقویتی یک سازوکار مؤثر در ارتقاء کارایی شبکه‌های عصبی *GAN* در تولید تصاویر با وضوح مکانی بالاتر باشد.

نتایج تجربی کسب‌شده از اجرای روش پیشنهادی حاکی از آن است که به‌کارگیری 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 به‌طور مستقیم به کیفیت آموزش شبکه وابسته است. استفاده از شبکه‌های عصبی آموزش‌یافته به کمک مناطق جغرافیایی دیگر احتمال تولید محتوای تصویری کاذب را افزایش داده و نمی‌توان از شبکه‌های آموزش‌یافته (به شکل تقویتی) در یک منطقه، انتظار تولید تصاویر مطلوب را در منطقه دیگر داشت. بدیهی است که وقوع تغییرات مکانی شدید ناشی از اختلاف‌زمانی تصاویر آموزشی مستخرج از سامانه گوگل-ارث و تصاویر چندطیفی مورداستفاده نیز می‌تواند احتمال ظهور پدیده‌های کاذب در روند افزایش وضوح مکانی را افزایش دهد. از این رو، تأکید اصلی این روش در به‌کارگیری تصاویر سامانه گوگل-ارث در کوتاه‌ترین اختلاف‌زمانی نسبت به تصاویر چندطیفی بوده تا کمترین اختلافات مکانی ممکن در داده‌های آموزشی نسبت به تصاویر چندطیفی وجود داشته باشند.

از دیگر نتایج تجربی کسب‌شده، ضرورت به‌کارگیری روش تناظریابی هیستوگرام میان تصاویر آموزشی و تصاویر چندطیفی در آماده‌سازی داده‌های آموزشی است. عدم اجرای این اقدام تأثیر مستقیم در کیفیت نتایج تولید تصاویر با وضوح مکانی بهتر داشته و احتمال تولید محتوای رادیومتریکی نادرست در مورد پدیده‌های تاریک تصویر را افزایش می‌دهد. به‌طور مثال، در زمان عدم به‌کارگیری روش تناظریابی هیستوگرام، احتمال بروز اختلاط در محتوای رادیومتریکی پدیده‌های آب و پوشش گیاهی مشاهده‌شده که این اتفاق در زمان به‌کارگیری روش تناظریابی هیستوگرام در تولید داده‌های آموزشی به شکل محسوسی کاهش‌یافته است.

به‌طورکلی 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 روئیت 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 مورداستفاده قرار گیرد.

۵- نتیجه‌گیری

در این مقاله راهکاری دومرحله‌ای به‌منظور بهبود وضوح مکانی تصاویر چند طیفی از طریق تولید تصاویر به کمک سوپررزولوشن و سپس تلفیق آن با تصاویر چند طیفی پیشنهاد شد. در تولید تصاویر با وضوح مکانی بهتر از یک شبکه عصبی تخاصمی از پیش آموزش‌یافته

[Downloaded from jgit.kntu.ac.ir on 2026-01-01]
[DOR: 20.1001.1.20089635.1402.11.2.3.6]
[DOI: 10.61186/jgit.11.2.41]

استفاده‌شده که یادگیری تقویتی آن بـه کمـک تصـاویر محلی مستخرج از سامانه گوگل-ارث اتفاق افتاده است. محـدودیت‌هـای فناورانـه بـه دلیـل فقـدان داده‌هـای آموزشی موردنظر، با استفاده از سـامانه گوگـل-ارث در تولید محتوا برای آموزش تقـویتی جبـران شـده اسـت. فقدان محتوای طیفی قابل‌مقایسه با تصاویر چند طیفی امکـان بهبـود هـم‌زمـان وضـوح مکـانی و طیفـی را در شبکه‌های عصبی فراهم نیاورده و ازاین‌رو در گام دوم از روش پیشنهادی، تصاویر تولیدشده با وضوح مکانی بهتر با تصاویر چند طیفی تلفیق‌شده‌اند. این موضـوع ضـمن حفظ محتوای طیفی، منجر به تولید تصاویری با وضـوح مکانی بهتر از تصـاویر چنـد طیفـی اولیـه شـده اسـت. توسـعه راهکـاری بـه‌منظـور بازیـابی محتـوای طیفـی هم‌زمان با بهبود وضوح مکـانی در یـک سـامانه هـوش مصنوعی و روش‌های یـادگیری عمیـق افـق تحقیقـاتی پیش‌رو در این حوزه محسوب می‌شود. مهم‌ترین چالش در این حوزه، موضوع هم مرجع سازی تصاویر بوده کـه تولیـد تصـاویر قـائم واقعـی و بـه‌کـارگیری روش‌هـای تناظریابی دقیق (همانند تناظریابی کمترین مربعـات بـا توابع هندسی منعطف) محور اجرای آن خواهـد بـود. از دیگر پیشنهادهای مـرتبط بـا ایـن حـوزه مـی‌تـوان بـه بکـارگیری ایـن روش در زمـان وقـوع حـوادث طبیعـی (سیل، زلزله، سونامی، زمین لغـزش) بـه منظـور تولیـد سریع تصاویر و با وضوح مکانی بالا جهت ارزیـابی اولیـه از میـزان خسـارت‌هـای وارده اشـاره کـرد. در چنـین شرایطی دسترسی سریع به تصاویر با وضوح مکانی بـالا یک چالش فنی تلقی شده که چنین راهکارهایی امکـان تامین اطلاعـات مناسـب را در شـرایط بحرانـی فـراهم می‌آورند. ارزیابی کارایی چنین راهکارهایی نیز می‌توانـد در دستورکار تحقیقات آتی قرار گیرد.

قدردانی و تشکر

نویسـندگان مقالـه از سـازمان جغرافیـایی و نیروهـای مسلح به دلیل در اختیار قـرار دادن تصـاویر مـاهواره‌ای کارتوست (*P5*) کمال تشکر خود را اعلان می‌دارند.

مراجع

[1] *F. Alidoost, M. Mobasheri, A. Abkar, "An efficient method to increase the spectral and spatial resolution in satellite images in urban environment", Iranian journal of Remote Sensing & GIS, vol. 4, no. 4, 2013.*

[2] *S. M . N. Niazi, M. Mokhtar Zade, F. Saeed Zadeh, "A Novel IHS-GA Fusion Method Based on Enhancement Vegetated Area", Journal of Geomatics Science and Technology, vol. 6, no. 1, pp. 235-248, 2016.*

[3] *Y. Ling, M. Ehlers, E. L. Usery, 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.*

[4] *Y. Zhang, 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.*

[5] *M. Bashirpoor, M. Valadan, Y. Maghsoodi, "Comparison of integration methods based on PCA and IHS in order to integrate Hyperion and Cartosat-1 images", Iranian journal of Remote Sensing & GIS, vol. 8, no. 4, pp. 17-30, 2016.*

[6] *R. Shahhoseini, T. S. Seyyedi, R. Habiballahi, "Comparative Evaluation of Image Fusion Methods for Hyperspectral and Panchromatic Data Fusion in Agricultural and Urban Areas", Geospatial Engineering Journal. 10, no. 2, pp. 63-78, 2019.*

[7] *M. Kabolizade, K. Rangzan, S. Mohammadi, "Application of fusion in satellite images the Landsat-8 and Sentinel-2 in environmental monitoring", Journal*

[Downloaded from jgit.kntu.ac.ir on 2026-01-01]
[DOR: 20.1001.1.20089635.1402.11.2.3.6]
[DOI: 10.61186/jgit.11.2.41]

applied RS and GIS techniques in natural resource, vol. 9, no. 3, pp. 53-71, 2018.

[8] K. Yaghoubi, A. R. Safdarinezhad, M. Jafari, "A method for determining the optimum parameter of the soft filters to image fusion in the frequency domain", Journal of Space Science and Technology, vol. 14, no. 3, pp. 23-37, 2021.

[9] M. Gargiulo, A. Mazza, R. Gaetano, G. Ruello, G. Scarpa, "Fast super-resolution of 20 m sentinel-2 bands using convolutional neural networks", Remote Sensing, vol. 11, no. 22, p. 26-35, 2019.

[10] S. Kalantari, M. J. Abdollahifard, S. Ahmadi, "Image Super-Resolution Using Analytical Edge Model", Journal of Iranian Association of Electrical and Electronics Engineers, vol. 15, no. 2, pp. 45-54, 2018.

[11] Z. Wei, K.-K. Ma, "Contrast-guided image interpolation", IEEE Transactions on Image Processing, vol. 22, no. 11, pp. 4271-4285, 2013.

[12] K. Jia, X. Wang , 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.

[13] C. Ledig et al., "Photo-realistic single image super-resolution using a generative adversarial network", presented at IEEE conference on computer vision and pattern recognition, Honolulu, HI, USA, 2017.

[14] C. Dong, C. C. Loy, K. He, X. Tang, "Learning a deep convolutional network for image super-resolution", presented at 13th European Conference on Computer Vision, Zurich, Switzerland, 2014.

[15] K. Zhang, J. Liang, L. Van Gool, R. Timofte, "Designing a practical degradation model for deep blind image super-resolution", presented at the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 2021.

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

[17] J. Kim, J. K. Lee, K. M. Lee, "Deeply-recursive convolutional network for image super-resolution", presented at the IEEE conference on computer vision and pattern recognition, Las Vegas, NV, USA, 2016.

[18] J.-B. Huang, A. Singh, N. Ahuja, "Single image super-resolution from transformed self-exemplars", presented at the IEEE conference on computer vision and pattern recognition, Boston, Massachusetts, USA, 2015.

[19] W. T .Freeman, T. R. Jones, E. C. Pasztor, "Example-based super-resolution", IEEE Computer graphics and Applications, vol. 22, no. 2, pp. 56-65, 2002.

[20] D. Glasner, S. Bagon, M. Irani, "Super-resolution from a single image", presented at the 2009 IEEE 12th international conference on computer vision, Kyoto, Japan, 2009.

[21] C. Dong, C. C. Loy, K. He, 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.

[22] M. R. U. Hoque, R. Burks, C. Kwan, J. Li, "Deep learning for remote sensing image super-resolution", presented at the 2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), NY, USA, 2019.

[23] M. Habibi, A. R. Ahmadyfard, 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.

[24] Y. Li, B. Li, "Super-Resolution of Sentinel-2

[Downloaded from jgit.kntu.ac.ir on 2026-01-01]
[DOR: 20.1001.1.20089635.1402.11.2.3.6]
[DOI: 10.61186/jgit.11.2.41]

Images at 10m Resolution without Reference Images", Preprints.org 2021, 2021040556. https://doi.org/10.20944/preprints202104.0556.v1

[25]X. Wang, L. Xie, C. Dong, Y. Shan, "Real-esrgan: Training real-world blind super-resolution with pure synthetic data", presented at the IEEE/CVF International Conference on Computer Vision, Montreal, Canada, 2021.

[26]X. Wang et al., "Esrgan: Enhanced super-resolution generative adversarial networks", presented at the European conference on computer vision (ECCV) workshops, Munich, Germany, 2018.

[27]Y. Zhu, S. Newsam, "Densenet for dense flow", presented at the 2017 IEEE international conference on image processing (ICIP), Beijing, China, 2017.

[28]S. Bharati, P. Podder, M. Mondal, V. B. 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.

[29] E. H. Helmer, B. Ruefenacht, "Cloud-free satellite image mosaics with regression trees and histogram matching", Photogrammetric engineering and remote sensing, vol. 71, no. 9, pp. 1079-1089, 2005.

[30]Xu. Li, Y. Zhang, Y. Gao, S. Yue, "Using guided filtering to improve gram-schmidt based pansharpening method for GeoEye-1 satellite images", presented at 4th International Conference on Information Systems and Computing Technology, Shanghai, China,2016.

[31] T. Maurer" ,How to pan-sharpen images using the gram-schmidt pan-sharpen method–A recipe", presented at International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Hannover, Germany. 2013.

[32]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. 1-20, 2019.

[33]S.Talebi, A. Zarea, S. Sadeghian, 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.

[34]C. Keeratikasikorn, 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.

[Downloaded from jgit.kntu.ac.ir on 2026-01-01]
[DOR: 20.1001.1.20089635.1402.11.2.3.6]
[DOI: 10.61186/jgit.11.2.41]

K. N. TOOSI UNIVERSITY OF TECHNOLOGY
FACULTY OF GEODESY AND GEOMATICS ENGINEERING

Journal of Geospatial Information Technology
Vol.11, No.2, Summer 2023

Research Paper

Spatial resolution improvement of the multispectral satellite images through integration of image fusion and deep learning methods in the absence of high spatial resolution images

Hamid Amini Amirkolaee [1], Alireza Safdarinezhad [2*], Hamed Amini Amirkolaee [3]

1- MSc student of photogrammetry, Department of Geodesy and Surveying Engineering, Tafresh University
2- Assistant professor, Department of Geodesy and Surveying Engineering, Tafresh University
3- Ph.D, School of Surveying Engineering, University of Tehran, Tehran, Iran

Abstract

Improving the spatial resolution of the multispectral images is one of the common pre-processing steps in reaching the maximum potential of these uimages in remote sensing applications. The presence of images with higher spatial resolution along with the multispectral images allows the process of improving the 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 the spatial resolution. Using the generative adversarial network (GAN) is one of the effective methods in this field that requires 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 in order to be used for training the network. Therefore, in this research, an approach with two main steps is designed to improve the spatial resolution of the 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 the 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 using the traditional pan-sharpening method. The results demonstrated that the proposed approach improved the spatial resolution of the multispectral images by 32.85% better than the best comparative method and maintained the spectral content without the need to provide extensive training data.

Key words: *Image fusion, Super-resolution, Multispectral images, Deep learning, Boost learning.*

Correspondence Address: *Department of Geodesy and Surveying Engineering, Tafresh University, Tafresh, Iran.*
Tel : *+989123258852*
Email: *safdarinezhad@tafreshu.ac.ir*

[Downloaded from jgit.kntu.ac.ir on 2026-01-01]
[DOR: 20.1001.1.20089635.1402.11.2.3.6]
[DOI: 10.61186/jgit.11.2.41]