

بهبود کیفیت تصاویر در ناوبری پهپاد با استفاده از روش فراتفکیک‌پذیری مبتنی بر شبکه عصبی کانولوشنی با نداشت چندلایه

مجید آقابابایی^۱، سید محمدرضا موسوی^۲، پیمان خزائی پول^۳، محمد خویشه^۴

majid_ghababaie@iust.ac.ir

۱- استادیار دانشکده مهندسی برق، دانشگاه علوم دریایی امام خمینی (ره)

۲- استاد دانشکده مهندسی برق، دانشگاه علم و صنعت ایران

۳- دانشجوی کارشناسی ارشد دانشکده مهندسی برق، دانشگاه علوم دریایی امام خمینی (ره)

۴- دانشجوی دکتری دانشکده مهندسی برق، دانشگاه علم و صنعت ایران

چکیده

ناوبری هوشمندانه و خودکار پهپاد براساس انطباق تصاویر دریافتی از پهپاد با تصاویر ماهواره‌ای یکی از جدیدترین گونه‌های ناوبری است که بسیار مورد توجه محققان و صنعت‌گران این حوزه قرار گرفته‌است. این روش هم از نظر جنگ الکترونیک و هم از نظر کارایی، زمانی موثر است که تصاویری با کیفیت بالا موجود باشد تا ویژگی‌های تصاویر را بتوان استخراج نمود. اما یکی از عواملی که سبب کاهش بهره‌ی استخراج ویژگی‌های تصاویر و انطباق آن با تصاویر ماهواره‌ای می‌شود، ماتی است. مات‌زدایی از تصاویر تبدیل به یک موضوع به‌روز و چالش برانگیز برای محققان شده است. در این مقاله روشی نوین برای بهسازی کیفیت تصاویر با استفاده از روش فراتفکیک‌پذیری مبتنی بر شبکه عصبی کانولوشنی CNN با چند لایه نداشت غیرخطی ارائه گردید که نقش بسزایی در مات‌زدایی و نویززدایی از تصاویر پهپاد دارد. نتایج نشان می‌دهد که این روش نسبت به روش‌های مطرح امروزی، کارایی بهتری دارد. به طوری که روش پیشنهادی میزان کیفیت تصویر را نسبت به روش‌های مطرح، براساس معیار نسبت پیک سیگنال به نویز (PSNR) تقریباً به اندازه ۵٪ افزایش می‌دهد.

واژگان کلیدی: ناوبری، سیستم موقعیت‌یاب جهانی، فراتفکیک‌پذیری، پراکندگی، شبکه عصبی کانولوشنی.

تاریخ دریافت مقاله : ۹۵/۰۴/۱۴

تاریخ پذیرش مقاله : ۹۵/۰۷/۲۸

۱- مقدمه

ناوبری علمی است که با آن می‌توان وسیله مورد نظر را از محلی به نام مبدا به محلی دیگر به نام مقصد هدایت نمود، به نحوی که ناوبر در هر زمان که نیاز باشد موقعیت خود را مشخص و در صورت نیاز گزارش نماید.

امروزه برای ناوبری کلیه تجهیزات نظامی مانند پهپادها، موشک‌ها و غیره از سیستم موقعیت‌یاب جهانی^۱ (GPS) استفاده می‌شود و با کمک خلبان خودکار آن‌ها را هدایت می‌کنند. علاوه بر این، روش‌های دیگری نیز وجود دارند که از سیگنال‌های GPS بی‌نیاز هستند، چرا که این سیگنال‌ها همیشه در دسترس نمی‌باشند و به همین دلیل، هنوز سازمان بین‌المللی هواپیمایی کشوری^۲ آن را به‌عنوان یک وسیله استاندارد و مطمئن برای مکان‌یابی وسایل پرنده نمی‌شناسد [۱]. یکی از بهترین روش‌های مستقل از سیگنال‌های GPS، برای هدایت خودکار وسایل پرنده از جمله پهپادها، انطباق تصاویر گرفته‌شده توسط دوربین پهپاد با تصاویر ماهواره‌ای است [۳و۲]. از آنجایی که تصاویر ماهواره‌ای از ارتفاع بسیار زیاد گرفته شده‌اند و مقیاس آن‌ها با تصاویر دریافتی از دوربین یک پهپاد که در نزدیکی زمین پرواز می‌کند تفاوت بسیاری دارد [۴]، لذا برای استخراج ویژگی‌های تصاویر و نگاشت یک به یک بین آن‌ها، نیاز به تصاویری عاری از هر گونه نویز و ماتی است. بنابراین برای تحقق این موضوع از روش‌های مطرح بهسازی استفاده می‌شود [۵].

در دو دهه گذشته، روش‌های مختلفی از قبیل روش‌های مبتنی بر فراتفکیک‌پذیری^۳ (SR) و همچنین روش‌های مبتنی بر کرنل در زمینه بهسازی تصاویر دریافتی از پهپادها استفاده گردیدند که در ادامه به معرفی روش‌های مطرح و مشکلات آن‌ها پرداخته می‌شود.

یکی از روش‌های مطرح شده در بعد نرم‌افزاری، جهت افزایش کیفیت تصویر چه از لحاظ تعداد پیکسل‌ها و چه از لحاظ کاهش مقدار نویز، روش فراتفکیک‌پذیری است. این روش از لحاظ نام‌گذاری، به دلیل آنکه قادر خواهد بود از محدوده توانایی سامانه تصویربرداری فراتر رود، فراتفکیک‌پذیری نامیده می‌شود.

هدف این روش‌ها، تخمین تصویر HR با استفاده از یک یا چندتصویر رزولوشن پایین^۴ (LR) است. به‌طور کلی می‌توان آن‌ها را به دو گروه تقسیم کرد.

- مبتنی بر تک‌تصویر [۵].

- مبتنی بر چندتصویر [۶].

روش‌های فراتفکیک‌پذیری مبتنی بر تک‌تصویر در زمینه بهسازی تصاویر پهپاد کاربرد بیشتری دارند که از جمله آن‌ها می‌توان به روش‌هایی مانند روش فراتفکیک‌پذیری مبتنی بر درون‌یابی، روش فراتفکیک‌پذیری مبتنی بر بخش‌بندی و دیگر روش‌ها اشاره کرد. هریک از این روش‌ها نیز دارای مشکلاتی می‌باشند که در ادامه به آن‌ها پرداخته می‌شود.

به منظور حل مشکل روش‌های مبتنی بر درون‌یابی ساده، از قبیل درون‌یابی^۵ خطی و مکعبی، تعدادی اصلاحات در آن‌ها ایجاد شده است. لی^۶ و همکارانش از یکسانی هندسی بین تصاویر LR و HR مبتنی بر واریانس محلی تصویر LR استفاده کردند. ژان^۷ و همکارانش یک الگوریتم درون‌یابی غیرخطی لبه با استفاده از فیلترهای تطبیقی و ترکیب داده‌ها را معرفی کردند. همچنین پژوهش‌گران یک روش مبتنی بر تکرار درون‌یابی با استفاده از دو مرحله ادراک و تصحیح پیکسل‌های تخمین را پیشنهاد کردند، اگر چه اصلاح شدن معایب روش درون‌یابی می‌تواند کیفیت تصویر را در حسگر (بهبود تیزسازی لبه و بینایی در یک درجه معین) بهبود ببخشد اما مصنوعات اساسی درون‌یابی مانند ماتی و نویز را نمی‌تواند کاملاً برطرف کند.

برای حل مشکلات مربوط به روش فراتفکیک‌پذیری مبتنی بر بخش‌بندی تصویر، فریمن^۸ و همکارانش نوعی از الگوریتم فراتفکیک‌پذیری مبتنی بر یادگیری را پیشنهاد کردند که با استفاده از مدل مارکوف، متناظر با بخش‌هایی تصویر LR ورودی از پایگاه داده خارجی، بخش‌های بهینه HR را تخمین می‌زند. یانگ^۹ و همکارانش نیز این روش را با استفاده از کدینگ پراکندگی^{۱۰} توسعه دادند [۷].

روش‌های مبتنی بر بخش‌بندی تصویر در مات‌زدایی و نویززدایی نسبت به روش‌ها فراتفکیک‌پذیری مبتنی بر درون‌یابی بهتر عمل می‌کنند، اگر چه بخش‌هایی از تصویر به دلیل تخمین نادرست از مولفه‌های فرکانس بالا ممکن است

⁶ Li

⁷ Zhang

⁸ Freeman

⁹ Yang

¹⁰ Sparse Coding

¹ Global Positioning System (GPS)

² International Civil Aviation Organization (ICAO)

³ Super-Resolution (SR)

⁴ Low Resolution (LR)

⁵ Interpolation

- ساختاری ساده و دقت خوبی نسبت به روش‌های فراتفکیک پذیری مبتنی بر یادگیری دارد.

- این روش دارای سرعت زیاد و کیفیت بسیار بالای بهسازی تصاویر است.

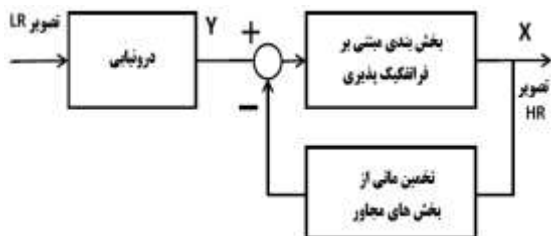
سازماندهی مقاله بدین صورت است که در بخش دوم به تشریح روش فراتفکیک‌پذیری مبتنی بر بخش‌بندی پرداخته شده است. در بخش سوم روش فراتفکیک‌پذیری مبتنی بر شبکه CNN با نداشت چندلایه تشریح می‌گردد. در بخش چهارم آموزش شبکه CNN توصیف می‌گردد. در بخش پنجم توضیح مختصری در مورد کرنل داده شده است. در بخش ششم نتایج شبیه‌سازی تشریح گردیده و در نهایت در بخش هفتم نتیجه‌گیری ارائه می‌گردد.

۲- روش فراتفکیک‌پذیری مبتنی بر بخش‌بندی

الگوریتم فراتفکیک‌پذیری روشی برای بازیابی جزئیات فرکانس بالا و حذف عوامل تخریب تصویر است. رابطه بین تصویر LR و HR در روش فراتفکیک‌پذیری مبتنی بر بخش‌بندی به صورت رابطه (۱) است.

$$Y = SHX + w \quad (1)$$

که در آن، Y مبین تصویر LR، H بیانگر عملگر ماتی، S مبین نمونه‌برداری کاهشی و X نشان‌دهنده تصویر HR است. پارامتر w مبین ماتی حاصل از بخش‌های مجاور بخش مورد نظر است که بیشترین تاثیر را در عملکرد مات‌زدایی روش فراتفکیک‌پذیری مبتنی بر بخش‌بندی دارد اما در بسیاری از الگوریتم‌ها نادیده گرفته می‌شود و به همین دلیل تخمین‌های نادرستی از بخش‌های تصویر LR ورودی بدست می‌آید، در نتیجه جزئیات فرکانس بالای تصویر به درستی بازیابی و بهینه نمی‌گردد. ساختار روش فراتفکیک‌پذیری مبتنی بر بخش‌بندی در شکل (۱) ارائه گردید.



شکل (۱) بلوک دیاگرام روش فراتفکیک‌پذیری مبتنی بر بخش‌بندی.

غیربهینه باقی بماند [۸]. در حالت کلی، روش‌های فراتفکیک‌پذیری مبتنی بر تک‌تصویر از هموارسازی لبه‌ها (استخراج ویژگی‌های لبه تصویر) و ماتی رنج می‌برند.

همچنین از روش‌های مورد استفاده در بهسازی پهناد می‌توان به روش‌های مبتنی بر کرنل مانند رگرسیون کرنل و کرنل وقفی نام برد [۹-۱۱] که مشکل همه آن‌ها در نگاشت و ثبت ویژگی‌ها تصاویر است. این روش‌ها بیشتر در زمینه حذف نویزهای قوی و ماتی کم استفاده می‌گردند.

هدف این مقاله بهبود عملکرد روش فراتفکیک‌پذیری مبتنی بر بخش‌بندی است که با استفاده از شبکه عصبی کانولوشنی^۱ (CNN) متشکل از چندین لایه نگاشت غیر خطی، اثرات (ماتی) نامطلوب بخش‌های مجاور بخش مورد نظر از تصویر مات شده، کاهش می‌یابد.

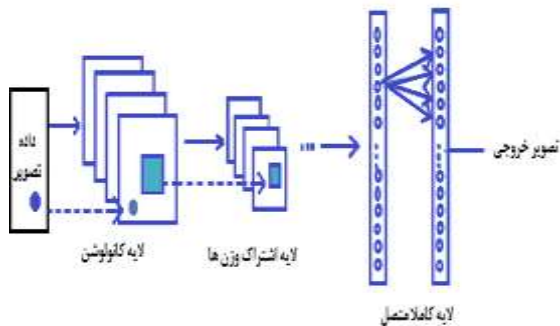
شبکه‌های عصبی CNN تا حد بسیار زیادی شبیه شبکه‌های عصبی مصنوعی می‌باشند. این نوع از شبکه‌ها متشکل از نرون‌هایی با وزن‌ها و بایاس‌های قابل یادگیری هستند. هر نرون تعدادی ورودی دریافت کرده و سپس حاصل ضرب وزن‌ها در ورودی را محاسبه می‌کند و در انتها با استفاده از یک تابع تبدیل (فعال‌سازی) غیرخطی نتیجه‌ای را ارائه می‌دهد [۱۲].

از تفاوت‌های شبکه عصبی CNN با شبکه‌های عصبی معمولی می‌توان به این نکته اشاره کرد که شبکه‌های عصبی معمولی یک ورودی در قالب بردار دریافت می‌کنند و سپس آن را از تعدادی لایه مخفی عبور می‌دهند و نهایتاً یک خروجی که نتیجه پردازش لایه‌های مخفی است در لایه خروجی شبکه ظاهر می‌شود. هر لایه مخفی از تعدادی نرون تشکیل شده که این نرون‌ها به تمام نرون‌های لایه قبل از خود متصل می‌شوند. نرون‌های هر لایه به صورت مستقل عمل کرده و هیچ ارتباطی با یکدیگر ندارند و آخرین لایه کاملاً متصل به لایه خروجی است. اگر از این ساختار برای تصاویر با پیکسل‌های زیاد استفاده شود، به تعداد بسیار زیادی نرون نیاز است، پس تعداد پارامترها به سرعت افزایش پیدا کرده و مشخص است اتصال کامل باعث اتلاف می‌شود. اما شبکه عصبی CNN از لایه‌های مجزا در امتداد عمق تصویر تشکیل شده است [۱۳].

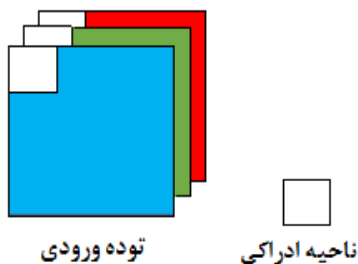
این روش چندین ویژگی مهم دارد:

¹ Convolutional Neural Network (CNN)

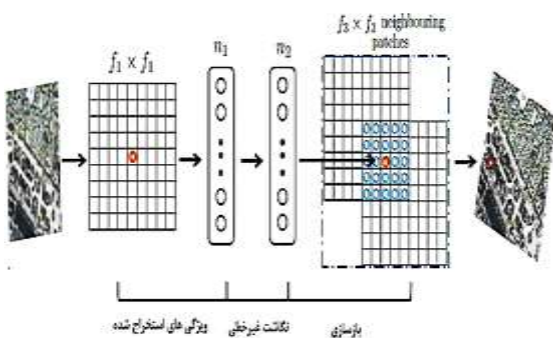
می‌شود. میزان وسعت این ناحیه کوچک که به آن ناحیه ادراکی گفته می‌شود، در امتداد محور عمق همیشه مساوی با تمام عمق ورودی است. شکل (۳) بیانگر این مسئله است. در این مقاله از تصویر ۸ بیتی خاکستری برای روش پیشنهادی استفاده گردیده است. تصویر LR دریافتی از پهپاد، Y نامیده می‌شود. همان‌گونه که در شکل (۴) نشان داده شده است. روش پیشنهادی برای اینکه اثر محدودیت بخش‌های مجاور در بخش مورد نظر از تصویر ورودی را حذف نماید (حذف مقدار w در روش فراتفکیک‌پذیری مبتنی بر بخش‌بندی)، شامل سه مرحله است. مرحله اول در این روش، استخراج ویژگی‌های تصاویر و ارائه دادن آن‌ها است.



شکل (۲) لایه‌های شبکه عصبی CNN.



شکل (۳) ارتباط ناحیه ادراکی در راستای عمق یک توده.



شکل (۴) ساختار روش فراتفکیک‌پذیری مبتنی بر شبکه CNN.

یکی از روش‌های مهم در زمینه بهبود روش فراتفکیک‌پذیری مبتنی بر بخش‌بندی، استفاده از کدینگ پراکندگی است [۱۴]. در روش مذکور برای تخمین نگاشت بین بخش‌های تصویر LR و HR از دو مجموعه دادگان استفاده می‌شود. در نهایت نیز بخش‌های استخراج شده LR را از فیلتر بالاگذر عبور می‌دهد تا جزئیات تصویر را استخراج نماید و از آن‌ها در زمینه بهسازی تصویر ورودی استفاده نماید. اما گاهی به دلیل تخمین نادرست بخش‌های تصویر ورودی، بعضی از جزئیات فرکانس بالا استخراج نمی‌گردد و سبب محدودیت‌هایی در آشکارسازی لبه‌های (عملکرد تیزسازی لبه) تصویر می‌شود در نتیجه تصویر حاصله دارای ماتی می‌گردد [۸].

در این مقاله، برای بهبود عملکرد روش فراتفکیک‌پذیری مبتنی بر بخش‌بندی در زمینه تخمین ماتی از بخش‌های مجاور، از شبکه عصبی CNN با نگاشت چندلایه استفاده گردیده است. زیرا روش پیشنهادی از یک مجموعه دادگان استفاده می‌کند، در نتیجه نسبت به روش کدینگ پراکندگی، اطلاعات کمتری از تصویر ورودی را از دست می‌دهد.

۳- روش فراتفکیک‌پذیری مبتنی بر شبکه عصبی کانولوشنی با نگاشت چندلایه

شبکه‌های CNN در دهه‌های قبل و اخیراً در طبقه‌بندی تصاویر به موفقیت بسیار بزرگی دست یافته‌اند [۱۳]. آن‌ها همچنین در زمینه‌های دیگر ماشین‌بینایی از قبیل آشکارسازی اشیاء و تشخیص چهره به کار برده شده‌اند.

شبکه‌های عصبی CNN از این واقعیت که ورودی شامل تصاویر است، استفاده کرده و معماری شبکه را به روش معقولی محدود می‌کنند [۱۲]. در معماری شبکه CNN معمولاً از سه نوع لایه استفاده می‌شود که این لایه‌ها عبارت‌اند از: لایه کانولوشن، لایه اشتراک ۱ وزن‌ها و لایه کاملاً متصل که دقیقاً همانند لایه‌های موجود در شبکه‌های عصبی معمولی است. شکل (۲) ساختار یک شبکه عصبی CNN را نشان می‌دهد.

در شبکه عصبی CNN زمانی که ورودی دارای ابعاد بالا (مانند تصاویر) باشد، اتصال نرون‌ها به تمام نرون‌های قبل از خود (در توده قبلی) غیرعملی است. بنابراین به جای اتصال کامل، هر نرون تنها به ناحیه کوچکی از توده ورودی متصل

¹ Pooling Layer

در این مقاله به بررسی مرحله نگاشت غیرخطی شبکه عصبی CNN پرداخته شد. زیرا این قسمت نقش بسزایی در عملکرد روش فراتفکیک پذیری دارد. بنابراین مهمترین عوامل تاثیرگذار در مرحله نگاشت غیرخطی که همان تعداد فیلترها و تعداد لایه‌ها است، مورد بررسی قرار گرفت.

فیلترهای مختلفی در این لایه با اندازه 1×1 یا 3×3 و یا 5×5 آزمایش گردید. فیلتر با اندازه 5×5 و 3×3 نتایج بهتری نسبت به فیلتر 1×1 دارند. بعد از انتخاب فیلترهای مناسب برای مرحله نگاشت در شبکه عصبی CNN، میزان تاثیرپذیری تعداد لایه‌های این بخش از شبکه مورد بررسی قرار گرفت. در این بررسی تعداد لایه‌های قسمت نگاشت (n) را برابر ۲ یا ۳ و یا ۴ انتخاب شده است، پس از بررسی فیلتر 5×5 با این تعداد لایه‌ها در قسمت نگاشت شبکه، نتایج خوبی در بهبودی تصویر خروجی نسبت به استفاده از فیلتر 3×3 حاصل نگردید. در نهایت فیلتر کانولوشنی با اندازه 3×3 برای مرحله نگاشت روش پیشنهادی انتخاب گردید.

تعداد و اندازه فیلترها در لایه‌های نگاشت برابر است. بهترین مقادیر بهبودی تصویر هنگامی به دست می‌آید که ۳ لایه در قسمت نگاشت وجود داشته باشد. در اولین لایه کانولوشنی موجود در قسمت نگاشت غیرخطی، فیلتر کانولوشنی با گام ۴ اعمال می‌گردد تا همان نقش نمونه برداری کاهشی (لایه اشتراک وزن‌ها) را انجام دهد و اندازه ویژگی‌های نگاشت شده به لایه کانولوشنی بعدی را کاهش دهد. در دومین لایه کانولوشنی بخش نگاشت، فیلتر کانولوشنی با گام ۱ اعمال می‌شود. در لایه کانولوشنی سوم نیز از فیلتر (کرنل) درون‌یابی دو خطی استفاده می‌گردد که وظیفه نمونه برداری افزایشی از ویژگی‌های نگاشت شده را دارد. با استفاده از کرنل دوخطی می‌توان اندازه تصویر را افزایش و در نتیجه ویژگی‌های لبه‌های تصویر را بهتر بدست آورد.

در مرحله سوم نیز بازسازی تصویر انجام می‌شود. در این مرحله پیش‌بینی رویهم افتادگی قسمت‌های رزولوشن بالا برای تولید یک تصویر با کیفیت انجام می‌شود. در واقع در این مرحله میانگین همانند فیلتر از پیش تعریف شده در یک دستگاه نگاشت ویژگی‌ها در نظر گرفته می‌شود و موقعیت هر کدام برابر یک بردار مسطح، از یک قسمت رزولوشن بالا است [۱۵]. عملکرد لایه سوم شبکه عصبی CNN به صورت رابطه (۴) نشان داده می‌شود:

این عملکرد ویژگی‌های تصویر رزولوشن پایین Y را استخراج می‌کند و هر ویژگی را همانند یک بردار اندازه ارائه می‌دهد و این بردار یک دستگاه نگاشت را شامل می‌شود. یکی از معروف‌ترین روش‌ها برای بازسازی تصاویر، استخراج ویژگی‌ها و ارائه آن‌ها با یک دستگاه پیش آموزش مبتنی بر PCA، Haar، DCT و غیره است که این معادل با کانولوشن یک تصویر با دستگاه فیلترهاست که نقش مهمی در شبکه CNN دارد و توسط شبکه عصبی بهینه‌سازی می‌شود. عملکرد لایه اول در شبکه عصبی CNN به صورت رابطه (۲) بیان می‌شود:

$$F_1(Y) = \max(0, W_1 * Y + B_1) \quad (2)$$

که در آن، W_1 و B_1 فیلترها و بایاس‌ها را ارائه می‌دهند و W_1 به اندازه $9 \times 9 \times 64$ است. W_1 از ۶۴ کانولوشن در تصویر استفاده می‌کند و هر کانولوشن به اندازه کرنل 9×9 است. خروجی از ۶۴ خصوصیت نگاشت^۱ ترکیب می‌شود و B_1 یک بردار ۶۴ بعدی است که هر المان به فیلتر وابسته است و از واحد جبران‌ساز خطی^۲ (ReLU) در پاسخ فیلترها که در واقع رابطی بین قسمت نگاشت غیرخطی و قسمت استخراج ویژگی‌ها است، استفاده می‌شود. در واقع لایه ReLU بر روی تک تک نرون‌های یک تابع فعال‌سازی مثل $\max(0, x)$ آستانه‌گذاری را بر روی صفر انجام می‌دهد (یعنی مقادیر منفی را صفر در نظر می‌گیرد). این کار تغییری در اندازه توده از مرحله قبل ایجاد نمی‌کند.

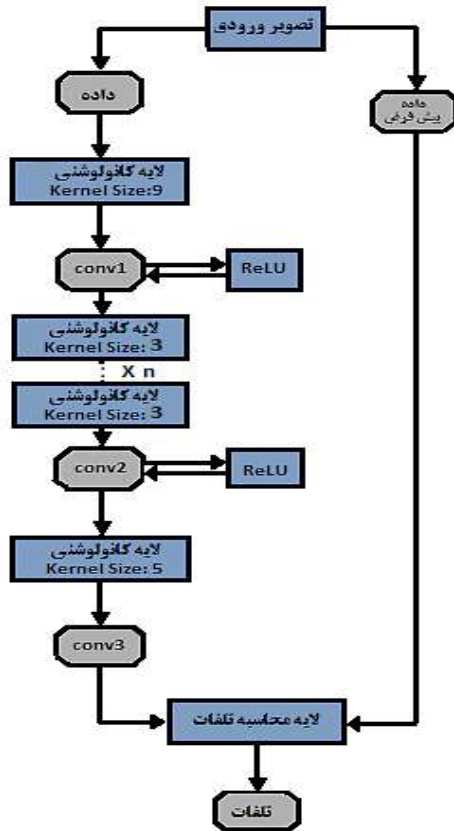
مرحله دوم، نگاشت غیرخطی ویژگی‌های به دست آمده از تصویر است. در این مرحله اولین لایه یک ویژگی ۶۴ بعدی برای هر قسمت استخراج می‌کند. در دومین عملگر هر یک از بردارهای ۶۴ بعدی به یک بردار ۳۲ بعدی نگاشت داده شده است. این معادل با به کار بردن ۳۲ فیلتری است که دارای پشتیبانی فضایی 1×1 است و این عملکرد در لایه دوم به صورت رابطه (۳) تعریف می‌گردد:

$$F_2(Y) = \max(0, W_2 * F_1(Y) + B_2) \quad (3)$$

که در آن، w_2 مبین اندازه‌ای برابر $1 \times 1 \times 64$ دارد و B_2 ، ۳۲ بعدی است. هر خروجی بردار ۳۲ بعدی یک قسمت تصویر رزولوشن بالا است که برای بازسازی استفاده خواهد شد. این امکان وجود دارد تا لایه‌های کانولوشنی بیشتری (با فضای پشتیبانی 1×1) برای افزایش غیرخطی استفاده شوند. اما این می‌تواند افزایش قابل توجهی در پیچیدگی مدل داشته باشد.

² Rectified Linear Unit (ReLU)

¹ Feature Map



شکل (۵) بلوک دیاگرام شبکه عصبی CNN.

۵- کرنل

روش فیلتر کردن با کرنل استرینگ^۳ (SK) روشی مقاوم در برابر نویز است که به وسیله تخمین گرادیان، ساختار محلی تصاویر را تعیین می‌کند و از این اطلاعات ساختاری برای تعیین اندازه کرنل استاندارد استفاده می‌نماید [۹]. شکل (۶) ساختار کرنل در امتداد لبه تصویر را نشان می‌دهد. یک تابع گرادیان محلی تخمینی از بین پیکسل‌های همسایه را اندازه می‌گیرد و از این تخمین برای وزن‌دهی اندازه‌گیری‌های مربوطه استفاده می‌شود. برای مثال، اگر یک پیکسل نزدیک لبه قرار گرفته باشد، پیکسل‌ها در همان سمت لبه اثر بیشتری دارند تا عمل فیلتر را انجام دهند. اگر شدت پیکسل‌ها براساس موقعیت مکانی آن‌ها بصورت $X_i = [x_i, y_i]$ در نظر گرفته شود، کرنل از رابطه (۹) محاسبه می‌گردد [۷]:

$$K(X_1 - X_i) = \sqrt{\det(c_i)} \exp\{-(X_1 - X_i)^T C_i (X_1 - X_i)\} \quad (۸)$$

^۳ Steering Kernel (SK)

$$F(Y) = W_3 * F_2(Y) + B_3 \quad (۴)$$

که در آن، W_3 برابر با ابعاد 5×5 و B_3 بردار ۱ بعدی است.

۴- آموزش شبکه عصبی CNN

در این مقاله از الگوریتم پس انتشار (BP) به منظور یادگیری نگاشت نقطه به نقطه تابع F ، نیازمند تخمین پارامترهای $\Theta = \{W_1, W_2, W_3, B_1, B_2, B_3\}$ است. این اقدام از طریق حداقل کردن تلفات یعنی بازسازی تصاویر $F(Y; \Theta)$ و همگرایی تصاویر رزولوشن بالای X دست‌یافتی است. یک دستگاه تصاویر رزولوشن بالا $\{X_i\}$ و متناظر آن‌ها تصاویر رزولوشن پایین $\{Y_i\}$ است. از متوسط مربع خطای^۱ (MSE) همانند تابع هزینه به صورت رابطه (۵) استفاده می‌شود:

$$L(\Theta) = 1/n \sum_{i=1}^n \|F(Y; \Theta) - X_i\|^2 \quad (۵)$$

که در آن، n بیانگر تعداد نمونه‌های تعلیم است. در این مقاله از ۱۰۰ تصویر LR دریافتی از پهباد استفاده گردیده است که اندازه هر تصویر 800×600 است. هر تصویر به ۲۰ بخش تقسیم می‌گردد. برای بهینه‌سازی وزن‌ها و بایاس‌های شبکه استفاده از MSE به عنوان تابع هزینه، برای دستیابی به بالاترین میزان نسبت پیک سیگنال به نویز^۲ (PSNR) است. PSNR یک معیار متریک برای ارزیابی کیفیت بازسازی تصاویر است و مرتبط به کیفیت ادراکی است. مقدار PSNR از رابطه (۶) بدست می‌آید:

$$PSNR = 10 \log_{10}(\text{blurred signal} / \text{MSE})[\text{dB}] \quad (۶)$$

وزن‌های جدید شبکه CNN توسط الگوریتم پس انتشار از روابط (۷) و (۸) به دست می‌آید.

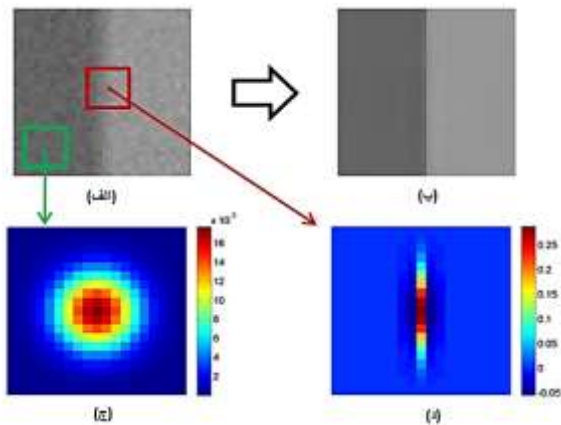
$$\Delta_{i+1} = 0.9 \cdot \Delta_i + \eta \frac{\partial L}{\partial \Delta_i} \quad (۷)$$

$$W_{i+1}^l = W_i^l + \Delta_{i+1} \quad (۸)$$

که در آن $i \in \{1, 2, 3\}$ ، تعداد لایه‌ها و تکرار آن‌ها را نشان می‌دهد و η نیز بیانگر نرخ یادگیری است. وزن فیلترها در هر لایه، به صورت تصادفی مقداردهی اولیه می‌شوند و از توزیع گوسی با میانگین صفر و انحراف استاندارد ۰/۰۰۱ (صفر برای بایاس‌ها) استفاده شده است. نرخ یادگیری برای همه لایه‌ها مقدار 10^{-4} انتخاب گردید. تابع هزینه MSE نیز فقط از اختلاف بین پیکسل‌های مرکزی تصویر X_i و تصویر ورودی شبکه محاسبه می‌شود. نحوه محاسبه تلفات توسط شبکه پیشنهادی در شکل (۵) ارائه گردیده است.

^۱ Mean Squared Error (MSE)

^۲ Peak Signal to Noise Ratio (PSNR)

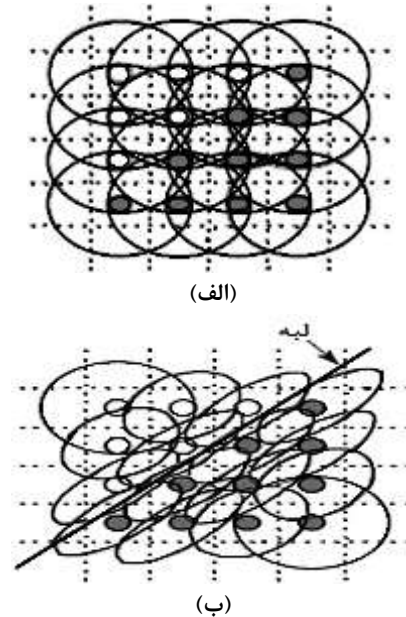


شکل (۷) مثالی از تیزسازی وقتی هندسی محلی: (الف) لبه دارای نویز و ماتی، (ب) لبه فیلتر شده و (ج) و (د) نویززدایی محلی وقتی / فیلتر تیزسازی [۹].

۶- بحث و شبیه سازی

در این مقاله، شبیه‌سازی شبکه عصبی CNN در نرم‌افزار Matlab 2015 و مدل Caffe بر روی سیستم با مشخصات Intel i7 CPU 4.0 GHZ انجام گردید. به منظور آزمایش کارایی شبکه عصبی CNN در روش فراتفکیک‌پذیری مبتنی بر بخش‌بندی، این روش را بر روی ۱۰۰ تصویر دریافتی از پهباد آزمایش کردیم. معیارهای PSNR، میزان تشابه تصویر پردازش شده و تصویر اصلی^۲ (SSIM) را در تصاویر دریافتی از پهباد محاسبه نموده و نتایج طبق جدول (۱) ارائه گردید. همان‌طور که در شکل (۱۰) ملاحظه می‌شود، در شبکه عصبی CNN با افزایش اندازه فیلترها در لایه‌های کانولوشنی، این نتایج حاصل شده است که هر چه اندازه فیلتر در لایه دوم (با شرط اینکه اندازه فیلتر لایه اول بیشتر از اندازه لایه سوم باشد)، بیشتر شود، تاثیر بهتر و کارآتری بر روی PSNR تصویر دارد.

همان‌طور که از شکل (۱۱) ملاحظه می‌شود، افزایش تعداد لایه‌ها در شبکه عصبی CNN (منظور از افزایش تعداد لایه‌ها در واقع اضافه کردن لایه‌های نگاشت غیرخطی است)، مناسب نمی‌باشد. اگر زمانی که شبکه دارای لایه‌های اشتراکی و کاملاً متصل به هم نباشد، بلکه شبکه عصبی CNN بسیار حساس به پارامترهای اولیه و نرخ یادگیری است. همان‌طور که در شکل (۱۲) ملاحظه می‌شود، از آنجایی که در ساختار پیشنهادی برای شبکه عصبی CNN، لایه‌هایی با نمونه‌برداری کاهشی و افزایشی وجود دارد، میزان PSNR افزایش می‌یابد.



شکل (۶) (الف) کرنل‌های استاندارد و (ب) کرنل‌های استرینگ در امتداد لبه [۹].

که در آن، X_i نشان دهنده یک مکان مشخص در پنجره کرنل فرمان به مرکزیت X_i است و C_i مبین یک ماتریس کوواریانس برآورد شده از مجموعه‌ای از شیب در یک پنجره تجزیه و تحلیل در اطراف X_i است.

یکی از روش‌های مطرح مبتنی بر کرنل برای مات‌زدایی تصاویر پهباد، روش تیزسازی وقتی هندسی محلی^۱ (GLAS) است که در شکل (۷) مراحل آن نشان داده شده است. این روش در مرحله اول مبتنی بر کرنل استرینگ برای مقاوم‌سازی در مقابل مصنوعات (ماتی و نیز) است و این هدف از ساختار محلی تصاویر دیجیتالی با تحلیل گرادین آن تصاویر تعیین می‌شود. بنابراین روش کرنل GLAS از رابطه (۱۰) به دست می‌آید:

$$S = K + qL \otimes K \quad (10)$$

که در آن، \otimes نشان دهنده عملگر کانولوشن، L بیانگر فیلتر لاپلاسی و q عددی مثبت که نشان دهنده میزان تیزسازی است و از به‌سازی تصویر بعد از تخمین رگرسیون محلی (میانگین وزن‌ها) با کرنل S به دست می‌آید. این روش فقط برای بهبود تصاویر با نویزهای قوی و ماتی کم به کار می‌رود و معیارهای اندازه‌گیری کیفیت در این روش بسیار کم هستند (برای توضیحات بیشتر در این زمینه به مرجع [۹] مراجعه شود).

² Structural Similarity Index Measure (SSIM)

¹ Geometric Locally Adaptive Sharpening (GLAS)

جدول (۱) نتایج دو معیار PSNR و SSIM در روش پیشنهادی.

PSNR [dB]	SSIM	تصاویر
۲۵/۴۷	۰/۷۹	تصویر شکل (۸-ب)
۲۶/۶۹	۰/۸۷	تصویر شکل (۸-ج)
۳۴/۲۸۶۹	۰/۶۵	تصویر شکل (۹-ب)
۳۶/۰۹۷	۰/۷۶	تصویر شکل (۹-ج)



(الف)



(ب)



(ج)

شکل (۹) (الف) تصویر هوایی اصلی از منطقه شهری، (ب) تصویر مات شده منطقه شهری و (ج) تصویر حاصل از روش فراتفکیک پذیری مبتنی بر شبکه CNN.



(الف)

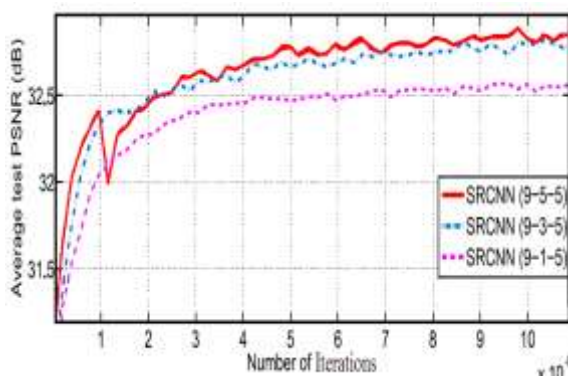


(ب)



(ج)

شکل (۸) (الف) تصویر هوایی اصلی از ناو جنگی، (ب) تصویر مات شده ناو جنگی و (ج) تصویر حاصل از روش فراتفکیک پذیری مبتنی بر شبکه CNN.



شکل (۱۰) نتایج PSNR حاصل از افزایش اندازه فیلتر در لایه میانی شبکه CNN.

همان طور که از جدول (۲) معلوم می‌گردد، روش پیشنهادی دارای PSNR و SSIM بالاتری نسبت به روش‌های مطرح فراتفکیک پذیری است.

با توجه به جدول (۱) و مقادیر PSNR و SSIM می‌توان نتیجه‌گیری کرد که روش فراتفکیک‌پذیری مبتنی بر شبکه عصبی CNN کیفیت تصاویر دریافتی از پهپاد را تقریباً به اندازه ۰.۵٪ افزایش می‌دهد که نشان‌دهنده افزایش کیفیت تصویر یا کاهش ماتی در تصویر با کیفیت پایین است. به علاوه، این روش تشابه ساختاری تصویر خروجی به تصویر اصلی را افزایش می‌دهد. همچنین نتایج مقایسه روش پیشنهادی فراتفکیک‌پذیری مبتنی بر شبکه CNN با روش‌های فراتفکیک‌پذیری مبتنی بر درون‌یابی، یادگیری، بخش‌بندی و رگرسیون کرنل و نیز روش کرنل و فقی هندسی (GLAS) مطابق جدول (۲) ارائه گردید.



(ب) (الف)



(د) (ج)



(و) (ه)

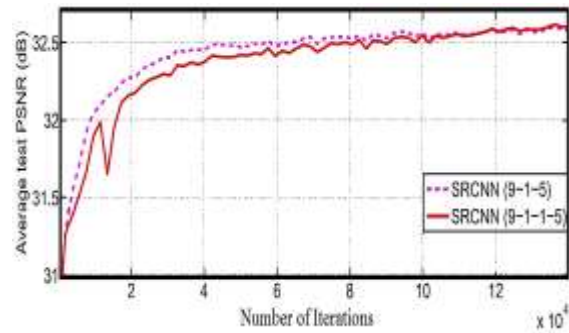


(ط)

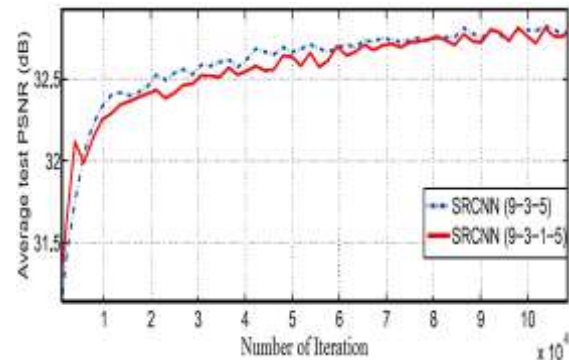
شکل (۱۳) (الف) تصاویر مات شده، تصاویر حاصل از روش فراتفکیک پذیری مبتنی بر (ب) درون یابی، (ج) یادگیری، (د) بخش بندی، (ه) و رگرسیون کرنل، (و) تصویر حاصل از روش کرنل افقی هندسی محلی و (ط) تصویر حاصل از روش فراتفکیک پذیری مبتنی بر شبکه CNN.

این نتایج نشان می دهد که روش پیشنهادی توانسته است مشکل اکثر روش های فراتفکیک پذیری که همان هموارسازی لبه (که مهمترین نقش را در ماتی تصویر دارد) را بر طرف نماید و ساختار تصویر را حفظ نماید.

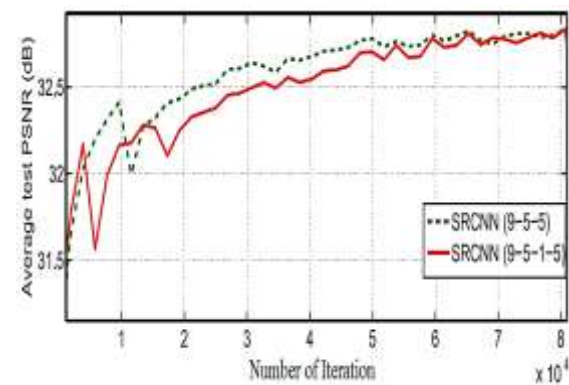
از مقایسه روش پیشنهادی با روش مطرح کرنل GLAS این نتیجه حاصل می گردد که روش کرنل GLAS به دلیل اینکه مبتنی بر کرنل استرینگ و کرنل گوسی است و این دو



(الف)

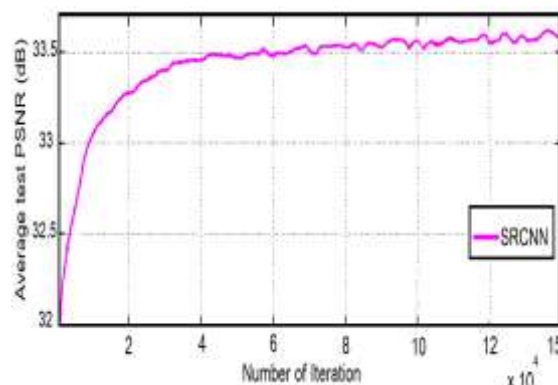


(ب)



(ج)

شکل (۱۱) نمودار مقایسه همگرایی شبکه های سه و چهار لایه CNN برای مدل های مختلف فیلتری (الف) مدل فیلتری ۹-۱-۵، (ب) مدل فیلتری ۹-۳-۵ و (ج) مدل فیلتری ۹-۵-۵.



شکل (۱۲) نتایج PSNR حاصل از ساختار پیشنهادی شبکه CNN.

۸- مراجع

- [۱] سید محمدرضا موسوی، مریم معاضدی، محمدجواد رضایی و امیر طباطبایی، "مقابله با اختلال در گیرنده‌های GPS"، انتشارات دانشگاه علم و صنعت ایران، ۱۳۹۴.
- [2] Lee, J. N. and Kwak, K., "A Trends Analysis of Image Processing in Unmanned Aerial Vehicle", International J. of Computer, Control, Quantum and Information Engineering, Vol.8, No 2, pp.180-186, 2014.
- [3] Niethamer, U., Rothmund, S., Schwader, U., Zeman, J. and Joswig, M., "Open Source Image-Processing Tools for Low-Cost UAV-Based Landslide Investigations", International Conference on Electrical and Control Engineering, Zurich, pp.161-166, 2011.
- [4] Nagai, M., Chen, T., Shibasaki, R., Kumai, H. and Ahmed, A., "UAV-Borne 3-D Mapping System by Multisensor Integration", IEEE Trans. on Geoscience and Remote Sensing, Vol.47, No.3, pp.701-708, 2009.
- [5] Bevilacqua, M., Roumy, A., Guillemot, C. and Morel, M. L. A., "Low-Complexity Single-Image Super-Resolution based on Nonnegative Neighbor Embedding", IEEE Conference on Computer Vision and Pattern Recognition, French, pp.1-10, 2012.
- [6] Huang, D., and Liu, H., "A Short Survey of Image Super Resolution Algorithms", Journal of Computer Science Technology Updates, Vol.2, pp.19-29, 2015.
- [7] Kang, W., Yu, SH., Ko, S. and Paik, J., "Multisensor Super Resolution using Directionally-Adaptive Regularization for UAV Images", Journal on the Science and Technology of Sensors and Biosensors, Vol.15, No.2, pp.1-27, 2015.
- [8] Kim, A., Park, J. and LEE, B.-U., "Removing Boundary Effect of a Patch-Based Super-Resolution Algorithm", IEICE Transactions Inf. & Systems, Vol.98, No.4 pp.974-979, 2015.
- [9] Zhu, X. and Milanfar, P., "Restoration for Weakly Blurred and Strongly Noisy Images", IEEE Conference on Computer Vision, USA, pp.103-109, 2011.
- [10] Dai, D., Timofte, R. and Gool, L., "Jointly Optimized Regressors for Image Super-Resolution", IEEE Signal Processing Letters, Vol. 23, No.5, pp.102-106, 2015.
- [11] Takeda, H., Farsiu, S. and Milanfar, P., "Kernel Regression for Image Processing and Reconstruction", IEEE Transactions on Image Processing, Vol.16, No.2, pp.349-366, 2007.
- [12] Cui, Z., Chang, H., Shan, S., Zhong, B. and Chen, X., "Deep Network Cascade for Image Super-Resolution", IEEE Conference on Computer Vision and Pattern Recognition, England, pp.49-64, 2014.

کرنل نقشی مهمی در آشکارسازی لبه‌های تصویر دارند. اما در زمینه وضوح تصویر دچار ضعف است. در حالی که در روش پیشنهادی از کرنل‌های گوسی برای استخراج ویژگی‌ها از تصویر ورودی و آشکارسازی لبه استفاده شده است و سبب بهبود عملکرد شبکه عصبی CNN گردید و توانسته وضوح تصویر خروجی را بالا برده و ماتی را به مقدار قابل قبولی حذف نماید.

جدول (۲) نتایج معیار PSNR و SSIM در روش‌های مختلف مات‌زدایی تصاویر دریافتی از پهپاد.

SSIM	PSNR	روش‌های مطرح به‌سازی تصویر	تصاویر
۰/۶۹	۱۷/۷۵	فرا تفکیک‌پذیری مبتنی بر درون‌یابی	شکل (۱۲)
۰/۷۲	۱۸/۶۱	فرا تفکیک‌پذیری مبتنی بر یادگیری	
۰/۸۲	۲۱/۹۴	فرا تفکیک‌پذیری مبتنی بر بخش‌بندی	
۰/۸	۲۵/۶۱	فرا تفکیک‌پذیری مبتنی بر رگرسیون کرنل	
۰/۷۸	۱۸/۱۵	کرنل وفقی هندسی محلی	
۰/۹۴	۳۲/۳۳	فرا تفکیک‌پذیری مبتنی بر شبکه CNN	

۷- نتیجه‌گیری

در این مقاله روش فرا تفکیک‌پذیری مبتنی بر شبکه عصبی CNN با نگاهی چندلایه معرفی گردید. روش پیشنهادی می‌تواند نگاهی دقیق و بهینه بین ویژگی‌های تصویر رزولوشن پایین ورودی و تصویر رزولوشن بالای خروجی ایجاد کند که این عمل در فرآیندهای پیش و پس از پردازش تصویر ورودی انجام می‌شود.

به منظور بررسی کارایی روش پیشنهادی در مقایسه با روش‌های قبلی ارائه شده، نتایج آماری این روش با روش‌های مطرح مات‌زدایی تصاویر پهپاد، روش‌های مختلف فرا تفکیک‌پذیری و کرنل GLAS مقایسه گردید. نتایج حاصل نشان می‌دهد که روش فرا تفکیک‌پذیری مبتنی بر شبکه CNN دارای ساختاری سریع است، زیرا تمام مراحل پردازش تصاویر را در لایه‌های خود به صورت یکجا و بهینه دارا است که نقش بسزایی در انطباق تصاویر برای ناوبری پهپاد دارد.

-
- [13] Ouyang, W., Wang, X., Zeng, X., Qiu, S., Luo, P., Tian, Y. and Tang, X., "Deepid-Net: Deformable Deep Convolutional Neural Networks for Object Detection", IEEE Conference on Computer Vision and Pattern Recognition, Boston, pp.2403-2412, 2015.
- [14] Yang, J., Wang, Z., Lin, Z., Cohen, S. and Huang, T., "Coupled Dictionary Training for Image Super-Resolution", IEEE Transactions on Image Processing, Vol.21, No.11, pp.3467-3478, 2012.
- [15] Chen, X., Xiang S., Liu, C. L. and Pan, C. H., "Vehicle Detection in Satellite Images by Parallel Deep Convolutional Neural Networks", IEEE Conference on Pattern Recognition, China, pp. 181-186, 2013.