نشریه علمی مهندسی فناوری اطلاعات مکانی

سال سیزدهم، شماره اول، بهار ۱۴۰۴

Vol.13, No.1, Spring 2025

۷۷ -۹۳

مقاله پژوهشی

ارزیابی بهبود روشنایی تصاویر هوایی به کمک روش‌های یادگیری عمیق

حسین زاور[1]، رضا شاه حسینی[2*]

۱- دانشجوی دکتری گروه فتوگرامتری و سنجش‌ازدور، دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، دانشکدگان فنی دانشگاه تهران

۲- دانشیار گروه فتوگرامتری و سنجش‌ازدور، دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، دانشکدگان فنی دانشگاه تهران

تاریخ دریافت مقاله: ۱۴۰۳/۰۷/۱۹ تاریخ پذیرش مقاله: ۱۴۰۴/۰۳/۰۷

چکیده

عکس‌برداری به کمک پرنده‌های بدون سرنشین به‌منظور پایش و بررسی وضع موجود یکی از پرکاربردترین موارد استفاده در نقشه‌برداری می‌باشد اما علی‌رغم فواید بسیار بالای این کار مشکلاتی نیز به همراه دارد. تنظیم نادرست دوربین در زمان عکس‌برداری، شرایط نامساعد جوی و تغییر شرایط نوری، مهم‌ترین عواملی هستند که باعث کاهش کیفیت تصاویر می‌شوند. به‌طورکلی روش‌های ارائه‌شده برای بهبود روشنایی را می‌توان به دودسته سنتی مانند استفاده از هیستوگرام و روش‌های مدرن شامل شبکه‌های عصبی و یادگیری عمیق که امروزه توجه بیش‌تر محققین را به خود جلب کرده است طبقه‌بندی کرد. هدف این تحقیق، بررسی عملکرد روش‌های یادگیری عمیق در بهبود روشنایی تصاویر هوایی است که به دلیل کاهش جزییات موجود ازنظر بصری و به طبع آن از دست رفتن اطلاعات طیفی، نامناسب می‌باشند. این تصاویر برای تولید محصولات مکانی گوناگون مانند تصویر اورتو و مدل دیجیتالی سطح زمین استفاده می‌شوند و بهبود روشنایی آن‌ها و بازیابی اطلاعات طیفی تأثیر مستقیمی بر روی محصولات مکانی خواهد داشت. به همین دلیل، در این مقاله ۳ روش مختلف یادگیری عمیق برای بهبود روشنایی تصاویر زمینی که در تحقیقات گذشته بهترین نتایج را به همراه داشته‌اند، بررسی می‌شوند و بهترین روش بر اساس ۱۰ معیار سنجش روشنایی انتخاب می‌گردد. داده‌های مورد ارزیابی، تصاویر هوایی اخذ شده از دو منطقه مختلف است که به دلیل روشنایی نامناسب دارای مناطقی با کاهش زیاد جزییات ازنظر بصری و از دست رفتن اطلاعات طیفی می‌باشند. آشکارسازی عوارض پنهان‌شده در مناطق سایه و مناطقی با میزان نور بیش‌ازاندازه به کمک بازیابی اطلاعات طیفی از طریق افزایش اختلاف بین مقادیر دیجیتالی پیکسل‌ها در این مناطق صورت می‌گیرد. نتیجه به‌دست‌آمده از بهترین روش انتخاب‌شده برای دو مجموعه داده، بر اساس شاخص شباهت ساختاری که یکی از مهم‌ترین شاخص‌های ارزیابی در بین ۱۰ شاخص استفاده شده است به ترتیب برای منطقه مورد مطالعه اول و دوم برابر ۰/۹۲ و ۰/۹۶ است.

کلیدواژه‌ها : فتوگرامتری، ادغام، بهبود روشنایی، اورتوفتو، یادگیری عمیق.

* نویسنده مکاتبه‌کننده: گروه فتوگرامتری و سنجش‌ازدور، دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، پردیس دانشکده‌های فنی، دانشگاه تهران، تهران، ایران.

تلفن: ۰۲۱۶۱۱۱۴۵۲۷

Email: rshahosseini@ut.ac.ir

[Downloaded from jgit.kntu.ac.ir on 2026-01-29]

[DOI: 10.61882/jgit.13.1.77]

۱- مقدمه

تولید محصولات مکانی مانند اورتو فتو و مدل دیجیتالی سطح زمین یکی از پرکاربردترین استفاده‌های رایج از پرنده‌های بدون سرنشین است چراکه به کمک آن‌ها در کمترین زمان می‌توان مناطق وسیعی را عکس‌برداری کرد. همواره عکس‌برداری هوایی به دلیل تغییرات ناگهانی شرایط جوی و روشنایی با چالش مواجه است. علاوه بر آن خطای انسانی در تنظیم پارامترهای روشنایی می‌تواند بر کاهش کیفیت عکس‌برداری مؤثر باشد. یکی از بهترین روش‌ها برای بهبود روشنایی، استفاده از روش‌های مبتنی بر ادغام تصاویر چند نوردهی می‌باشد (*MEF*).[1] روش کلی *MEF* استفاده از تصاویر تولیدشده با کاهش و افزایش روشنایی تصویر اصلی بوده که با ادغام آن‌ها یک تصویر با روشنایی بهینه تولید می‌کند [۱].

امروزه با ساخت دوربین‌های عکس‌برداری با تکنولوژی بالا توانایی کاهش اثرات محیطی از منطقه تصویربرداری بسیار بهبودیافته است اما همچنان به‌طور کامل نتوانسته‌اند که این چالش را برطرف کنند. تغییرات شرایط جوی و نوری در تصویربرداری هوایی یک پدیده رایج می‌باشد که این تغییرات باعث کاهش کیفیت تصاویر در مناطقی از آن‌ها خواهد شد. به‌طور مثال وجود نور بیش‌ازحد یا کمبود نور، مانند سایه، امری رایج است و بهبود روشنایی در چنین تصاویری برای استفاده در تولید محصولات نقشه‌برداری به دلیل پنهان شدن برخی عوارض از اهمیت بالایی برخوردار است[۲]. با توجه به وجود مناطقی با تراکم بالای عوارض مانند مناطق شهری عدم روشنایی مناسب به دلایل مختلف همچون زاویه نامناسب دوربین و یا شرایط جوی ممکن است بسیاری از جزییات موجود درصحنه وضوح مناسبی را نداشته باشند به همین دلیل نیاز به بهبود روشنایی بسیار ضروری می‌باشد. یکی از بهترین روش‌های موجود، روش‌های مبتنی بر *MEF* می‌باشند که نتایج به‌دست‌آمده از آن‌ها اثبات‌کننده این موضوع است. روش‌های متعددی مبتنی بر ادغام چند نوردهی به دلیل استقبال بالا از این روش ارائه‌شده‌اند که به‌طورکلی می‌توان به دو حوزه مکان‌مبنا و انتقال-مبنا دسته‌بندی کرد. برخلاف روش‌های مبتنی بر حوزه مکان که در فضای موجود و بدون تغییر فضا انجام می‌شوند، روش‌های انتقال مبنا فضای موجود را به فضای دیگری تبدیل و پس از اعمال به فضای اولیه برمی‌گردانند[۳]. با توجه به این‌که تاکنون ارزیابی مناسبی از روش‌های بهبود روشنایی برای تصاویر هوایی انجام‌نشده است، در این تحقیق با استفاده از روش‌های بهبود روشنایی در حوزه یادگیری عمیق، میزان تأثیر هر یک از این روش‌ها در بهبود روشنایی این نوع تصاویر مورد ارزیابی قرار خواهد گرفت.

۲- پیشینه تحقیق

تاکنون مطالعات زیادی درباره روش‌های مبتنی بر *MEF* انجام شده است و تکنیک‌های توسعه‌یافته در هر دو حوزه سنتی و هوش مصنوعی توسعه‌یافته‌اند. برت و همکاران (۱۹۹۳) یکی از اولین مطالعات را روی *MEF* انجام دادند که یک روش سنتی مبتنی بر هرم برای انجام چندین کار ادغام تصویر ازجمله ادغام تصویر مرئی-مادون قرمز، ادغام تصویر چند فوکوس و ادغام تصویر چند نوردهی هست[۴]. ابتدا تصاویر منبع از طریق تبدیل هرمی تجزیه می‌شوند. سپس ضرایبی برای تشکیل یک هرم برای تصویر کوچک‌شده انتخاب می‌شوند. درنهایت، تصویر کوچک‌شده از طریق تبدیل هرم معکوس بازیابی می‌شود[۱]. تعدادی از روش‌های مرسوم به‌منظور بهبود روشنایی بر روی هیستوگرام تصاویر تغییراتی را انجام می‌دهند. بهبود روشنایی در سرتاسر یک عکس بدون توجه به توزیع مکانی آن‌ها، پردازش یکسانی را روی همه پیکسل‌های تصویر انجام می‌دهد. بااین‌حال، مناطق روشن ممکن است پس از بهبود خطی اشباع‌شده و باعث از بین رفتن جزئیات در نتایج بهبودیافته شوند. یکی از روش‌ها برای جلوگیری

[1] *Multi Exposure Fusion*

[Downloaded from jgit.kntu.ac.ir on 2026-01-29]
[DOI: 10.61882/jgit.13.1.77]

از اشباع، یکسان‌سازی هیستوگرام[1] *(HE)* [۵] بوده که می‌تواند کنتراست را به‌طور مؤثر بهبود بخشد و به یک تکنیک پرکاربرد تبدیل شود *[۲]*. بسیاری از مشتقات *HE* برای در نظر گرفتن برخی محدودیت‌ها مانند حفظ روشنایی و محدودیت کنتراست پیشنهاد شده‌اند[۹,۸,۷]. یکسان‌سازی هیستوگرام ممکن است باعث از دست دادن جزئیات در برخی مناطق محلی شوند، زیرا پردازش به‌صورت کلی نمی‌تواند تضمین کند که همه مناطق محلی به‌خوبی بهبودیافته‌اند، اما به‌طورکلی، با استفاده تکنیک‌های محلی می‌تواند به نتایج بهتری دست‌یافت. یکسان‌سازی هیستوگرام محلی استراتژی پنجره کشویی را برای انجام یکسان‌سازی هیستوگرام به صورت محلی اتخاذ می‌کند[۷]. لئو و همکاران(۲۰۱۹) یک روش *MEF (DSIFT_EF)* مبتنی بر تبدیل ویژگی ثابت مقیاس (*SIFT*) پیشنهاد می‌دهد[۹]. در این روش، ابتدا برای استخراج جزئیات محلی از تصاویر منبع از توصیفگر *SIFT* استفاده و سپس برای حذف مصنوعات متحرک درصحنه استفاده می‌شود. همچنین لی و همکاران(۲۰۱۸) *MEFA* را بر اساس وزن‌های تطبیقی پیشنهاد می‌دهند[۱۰]. به‌طور خاص، آن‌ها دو تابع وزن را تعریف می‌کنند که کیفیت پیکسل را به ترتیب در مورد روشنایی کلی و گرادیان‌ها وزن دهی می‌کند. وزن نهایی ترکیبی از این دو وزن است. تصاویر ورودی برای هر مجموعه شامل یک سری از تصاویر با درجات روشنایی کم به زیاد می‌باشند. در بهبود روشنایی تصاویر، روش‌های بهینه‌سازی نیز به کار گرفته‌شده‌اند.

در سال‌های اخیر روش‌های یادگیری عمیق درزمینه *MEF* به کار گرفته‌شده است. روش‌های ارائه‌شده برای بهبود روشنایی را می‌توان به دودسته کلی نظارت‌شده و نظارت‌نشده تقسیم‌بندی کرد. پرابهاکار و همکارانش (۲۰۱۷) روش *DeepFus* را ارائه کردند که به‌عنوان یکی از اولین روش‌های ارائه‌شده در این زمینه مطرح است[۱۱]. ابتدا تصاویر ورودی به داده‌های کانال رنگی *YCbCr* تبدیل می‌شوند. سپس یک شبکه عصبی کانولوشنی که از سه لایه ویژگی، ادغام و لایه بازسازی تشکیل‌شده است، برای ترکیب کانال *Y* استفاده می‌شود. کانال‌های کرومینانس (*Cb* و *Cr*) با استفاده از وزن‌های ادغام به‌دست‌آمده، ادغام می‌شوند. لی و ژانگ(۲۰۱۸) با استفاده از مدل‌های از پیش آموزش‌دیده مشکل کمبود داده آموزشی را برطرف می‌کنند[۱۲]. مدل‌های شبکه عصبی کانولوشنی(*CNN*)[2]آموزش‌دیده در کارهای دیگر برای استخراج ویژگی‌های تصاویر چند نوردهی استفاده می‌شود و سپس وزن‌های ادغام بر اساس این ویژگی‌ها محاسبه می‌گردد. این روش (*MEFCNN*) می‌تواند تصاویر چند منبع را ترکیب کند. ین و چن(۲۰۲۰) یک روش مبتنی بر رمزگذار-رمزگشا[3] برای ترکیب دو تصویر با نوردهی متفاوت پیشنهاد می‌دهند[۱۳]. نوآوری اصلی این روش این است که یک رمزگذار با هدایت محتوای قبلی (*CPG*)[4]و یک رمزگشا با هدایت پیش از جزئیات (*DPG*)[5]طراحی‌شده است تا بتواند تصاویر را در سطح پیکسل و سطح ویژگی ترکیب کند. شبکه‌های متخاصم مولد (*GAN*)[6]نیز برای حل مشکل *MEF* استفاده‌شده‌اند. چن و همکاران(۲۰۲۰) یک روش مبتنی بر *GAN* را برای انجام *MEF* پویا پیشنهاد می‌دهند[۱۴]. این روش از سه جزء اصلی تشکیل‌شده است: یک شبکه تخمین هموگرافی برای جبران حرکت دوربین، یک ژنراتور برای تولید تصویر ترکیبی و یک شبکه تشخیص‌دهنده برای یادگیری متخاصم. روش *MEF* نظارت‌شده از تغییر مقادیر شدت تصاویر معمولی برای ایجاد تصاویر حقیقت زمینی استفاده

[1] *Histogram Equalization*

[2] *convolutional neural network*

[3] *Encoder and Decoder*

[4] *content-prior guided*

[5] *detail prior guided*

[6] *Generative Adversarial Network*

[Downloaded from jgit.kntu.ac.ir on 2026-01-29]

[DOI: 10.61882/jgit.13.1.77]

می‌کند. این روش می‌تواند از مدل‌هایی استفاده کند که از قبل در دیگر کارها آموزش داده‌شده‌اند، یا از نتایج ادغام سایر روش‌های *MEF* به‌عنوان داده آموزشی استفاده نماید. همه این راه‌ها ممکن است مشکل کمبود داده‌های آموزشی واقعی را برطرف نکند، ازاین‌رو روش‌های دیگری بر اساس روش‌های نظارت‌نشده ارائه شدند.

یانگ و همکاران (۲۰۲۱) روش *MEF* بدون نظارت مبتنی بر شبکه مولد تخاصمی (*GAN*) را توسعه دادند[۱۵]. *GANFuse* یک روش بدون نظارت است که از یک تابع ضرر بدون نظارت استفاده می‌کند. به‌طور خاص، تابع از دست دادن برای اندازه‌گیری شباهت بین تصویر ترکیبی و تصاویر منبع به‌جای تصاویر واقعی طراحی‌شده است. *GANFuse* از یک ژنراتور و دو تشخیص‌دهنده تشکیل‌شده است. هر تمایز برای تشخیص تصویر ترکیب‌شده و یک تصویر منبع استفاده می‌شود. لیانگ و ما (۲۰۲۳) به دلیل عدم وجود داده‌های مناسب برای یادگیری چند نوردهی از روش یادگیری متضاد[1]برای بهبود خروجی نهایی ازنظر بافت و روشنایی استفاده می‌کنند[۱۶]. در این روش می‌توان برخلاف سایر روش‌های دیگر حتی از دو تصویر با میزان روشنایی بیش‌ازاندازه و یا روشنایی کم استفاده کرد. پس از اعمال یک الگوریتم یادگیری عمیق کانولوشن غیر نظارت‌شده و به دست آوردن ویژگی‌های تصاویر ورودی و ادغام نهایی آن‌ها در یک یادگیری متضاد میزان روشنایی تصویر نهایی را بر اساس آموزش انجام‌شده به میزان روشنایی معمول[2] تبدیل کند تا میزان بافت و روشنایی تصویر نهایی مناسب باشد. نمونه‌هایی از تصاویر با میزان روشنایی متعادل به‌عنوان نمونه‌های مثبت و تصاویر با میزان روشنایی کم و یا زیاد به‌عنوان نمونه‌های منفی در نظر گرفته می‌شود و شبکه یادگیری متضاد آموزش می‌بیند.

تاکنون روش‌های یادگیری عمیق متعددی به‌منظور بهبود روشنایی تصویر ارائه‌شده‌اند اما هیچ‌کدام از روش‌های ارائه‌شده بر روی تصاویر هوایی در مناطق مختلف و با شرایط نوری نامناسب آزموده نشده است. روش‌های مبتنی بر یادگیری عمیق نیازمند سخت‌افزارهای قوی بوده، به همین دلیل بسیاری از این روش‌ها به‌منظور آموزش شبکه و تست آن از تصاویر باکیفیت بسیار پایین استفاده می‌کنند. به دلیل محدودیت‌های موجود در روش‌های یادگیری عمیق، تمرکز بیشتر تحقیقات انجام‌شده بر روی‌داده‌های با حجم کم می‌باشد. همچنین به دلیل عدم وجود داده کافی برای موضوعات خاص، روش‌های نظارت‌نشده به‌عنوان محصول نهایی در تحقیقات گذشته ارائه‌شده است. اکثر تصاویر هوایی اخذشده دارای کیفیت بالا بوده و کاهش کیفیت آن‌ها به دلیل کاهش کیفیت در محصولات مکانی امکان‌پذیر نیست. در این مقاله سعی می‌شود با روشی ابتکاری در کنار حفظ کیفیت تصاویر هوایی و استفاده از تاثیر روش‌های یادگیری عمیق در بهبود روشنایی، توانایی این روش‌ها بر روی تصاویر هوایی مورد ارزیابی قرار گیرد. ارزیابی از عملکرد این نوع الگوریتم‌ها بر روی‌داده‌های حجیم‌تر و به‌طور خاص تصاویر هوایی می‌تواند پاسخ مناسبی به توانایی الگوریتم‌های عمیق برای موضوع موردنظر باشد. روش‌های یادگیری عمیق مانند روش‌ها سنتی نیاز به تعیین حد آستانه گذاری برای هر مجموعه تصاویر به‌صورت جداگانه ندارند و می‌توان برای تصاویر از صحنه‌های مختلف و بدون تعیین پارامترهای متعدد که در روش‌های سنتی وجود دارند، استفاده کرد.

۳- داده‌های مورد استفاده

داده‌های مورداستفاده در این تحقیق تصاویر هوایی اخذشده به کمک پهپاد فانتوم ۴ پرواز دو منطقه مختلف است که به دلیل شرایط نامناسب محیطی از کیفیت بصری مناسبی برخوردار نبوده و در مناطقی از آن‌ها اطلاعات موجود چه ازنظر بصری و چه ازنظر طیفی محسوس نیست. داده‌های تهیه‌شده را می‌توان

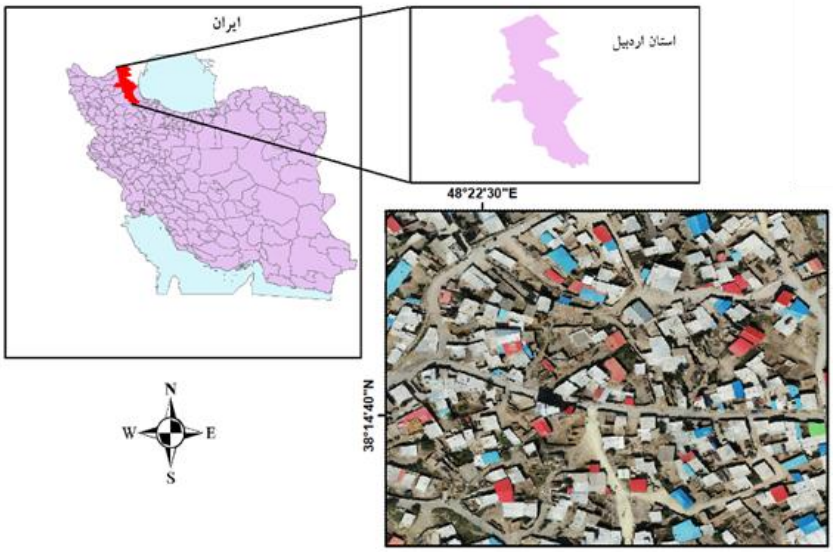
[1] *Contrastive Learning*
[2] *Normal*

[Downloaded from jgit.kntu.ac.ir on 2026-01-29]
[DOI: 10.61882/jgit.13.1.77]

ازنظر شرایط نوری به دودسته تقسیم کرد. اول، شرایطی که در آن نور محیط به دلیل وجود ابر و یا تنظیم نامناسب دوربین در زمان عکس‌برداری کم بوده که در این مناطق، عوارض داخل مناطق سایه به‌خوبی دیده نمی‌شوند. دوم، شرایطی که در آن میزان نور دریافتی به دلیل انعکاس بالای نور همچون سقف‌ها و ساختمان‌ها زیاد بوده و باعث می‌شود جزییات و عوارض موجود به‌خوبی دیده نشوند.

۳-۱- منطقه مطالعاتی اول

منطقه مورد مطالعاتی اول روستای آقا باقر در استان اردبیل است و شامل ۶۹۹ عکس هوایی اخذشده به کمک پهپاد فانتوم ۴ پرو بوده که اطلاعات نوع سنجنده در جدول(۱) و محدوده منطقه موردمطالعه در شکل(۱) آورده شده است. ارتفاع پرواز ۷۰ متر، میزان *iso* ۱۰۰ و سرعت شاتر ۱/۱۰۰۰ در نظر گرفته شده است. میزان روشنایی در بخش‌هایی از تصاویر مانند شکل(۲) که دارای سایه می‌باشند مناسب نبوده و عوارض داخل این محدوده‌ها به‌خوبی قابل‌تشخیص نمی‌باشند.



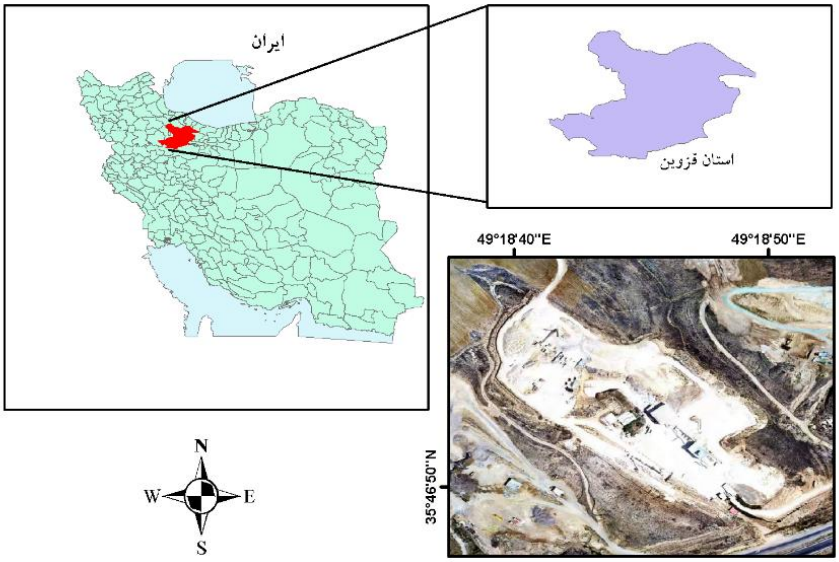
شکل۱: محدود منطقه مطالعاتی اول

شکل۲: نمونه تصویر برای منطقه مطالعاتی اول

[Downloaded from jgit.kntu.ac.ir on 2026-01-29]
[DOI: 10.61882/jgit.13.1.77]

۳-۲- منطقه مطالعاتی دوم

منطقه موردمطالعه دوم، واقع در استان قزوین، شامل ۱۲۴ عکس هوایی اخذشده به کمک پهپاد فانتوم ۴ پرو بوده که اطلاعات نوع سنجنده درجدول(۱) و محدوده مورد مطالعه در شکل(۳) آورده شده است. ارتفاع پرواز ۹۰ متر، میزان *iso* ۱۰۰ و سرعت شاتر ۱/۱۲۵۰ در نظر گرفت شده است. میزان روشنایی در بخش‌هایی از تصاویر به دلیل وجود موادی همچون آهک در سطح محدوده و بازتاب زیاد نور بسیار بالا بوده و جزییات موجود در قسمت‌هایی از تصاویر مانند شکل(۴) قابل مشاهده نمی‌باشند.



شکل۳: محدوده منطقه مطالعاتی دوم

شکل۴: نمونه تصویر برای منطقه مطالعاتی دوم

[Downloaded from jgit.kntu.ac.ir on 2026-01-29]
[DOI: 10.61882/jgit.13.1.77]

جدول ۱: مشخصات سنجنده

مشخصات	توضیحات
سایز سنسور	۱ اینچ
نوع سنسور	*CMOS*
سایز پیکسل	۲,۴ میکرون
تعداد پیکسل	۲۰ میلیون
زاویه دید	۸۴ درجه
لنز	۸,۸ میلی‌متر
محدوده ایزو (ویدئو)	۱۰۰ - ۳۲۰۰ (خودکار) ۱۰۰ - ۶۴۰۰ (دستی)
محدوده ایزو (عکس)	۱۰۰ - ۳۲۰۰ (خودکار) ۱۰۰ - ۱۲۸۰۰ (دستی)
سرعت شاتر مکانیکی	۸ - ۱/۲۰۰۰ ثانیه
سرعت شاتر الکترونیکی	۸ - ۱/۸۰۰۰ ثانیه

۴- روش مورد استفاده

برای بررسی بهبود روشنایی تصاویر هوایی از سه الگوریتم یادگیری عمیق *DPE*[1][۱۷] *MEFNET*[2] [۱۸]و *IFCNN*[3][۱۹] استفاده می‌شود. ژانگ(۲۰۲۱) جدیدترین روش‌های موجود در حوزه بهبود روشنایی را مورد ارزیابی قرار می‌دهد که روش‌های یادگیری عمیق نام‌برده شده، دارای نتایج خوبی بر اساس شاخص‌های روشنایی می‌باشند. علاوه بر این هرکدام از روش‌های ذکرشده در ارزیابی با چندین روش یادگیری عمیق عملکرد بهتری را به نمایش می‌گذارند[۱]. از طرفی ساختار الگوریتم‌های انتخابی قابلیت اجرا بر روی تصاویر پهپاد با توجه با محدودیت‌های سخت‌افزاری را دارا می‌باشند. بنابراین هرکدام از این الگوریتم‌ها بر روی‌داده‌های موجود با روشنایی نامناسب توانسته‌اند نتیجه‌های قابل قبولی ارائه دهند اما کارایی این شبکه‌ها بر روی تصاویر هوایی موردبررسی قرار نگرفته‌اند.

[1] *deep perceptual enhancement*

[2] *Deep Guided Learning for Fast Multi-Exposure Image Fusion*

[3] *image fusion framework based on convolutional neural network*

۴-۱- الگوریتم *MEF_ NET*

این الگوریتم از چند بخش مهم شامل درون‌یابی دوخطی[4]، نقشه وزن و فیلترهای راهنما[5] تشکیل‌شده است که ساختار کلی آن در شکل(۵) نشان داده‌شده است. در ابتدا تصاویر ورودی ازنظر اندازه کوچک‌شده و سپس وارد شبکه می‌شوند تا نقشه وزن با رزولوشن پایین تخمین زده شود. تعداد تصاویر ورودی در این الگوریتم اختیاری بوده و از دو تصویر و بیش‌تر را می‌توان به‌عنوان ورودی به آن معرفی کرد. حال با داشتن وزن‌های تصاویر با رزولوشن پایین و تصویر اصلی، نقشه وزن با رزولوشن بالا با استفاده از فیلترهای راهنما به دست می‌آید. درنهایت خروجی از مجموع وزن‌های بالا در تصاویر اصلی استخراج می‌شود. برای بررسی خروجی الگوریتم از شاخص *MEF_SSIM* استفاده می‌شود. برای آموزش شبکه، ۶۹۰ عکس از صحنه‌های مختلف که دارای روشنایی مختلف هستند استفاده‌شده است. ۹۰ عکس به‌منظور تست شبکه و ۶۰۰ عکس برای آموزش شبکه تقسیم‌بندی شدند. برای بهینه‌سازی از روش *Adam* و از تابع فعال‌ساز

[4] *BiLinear*

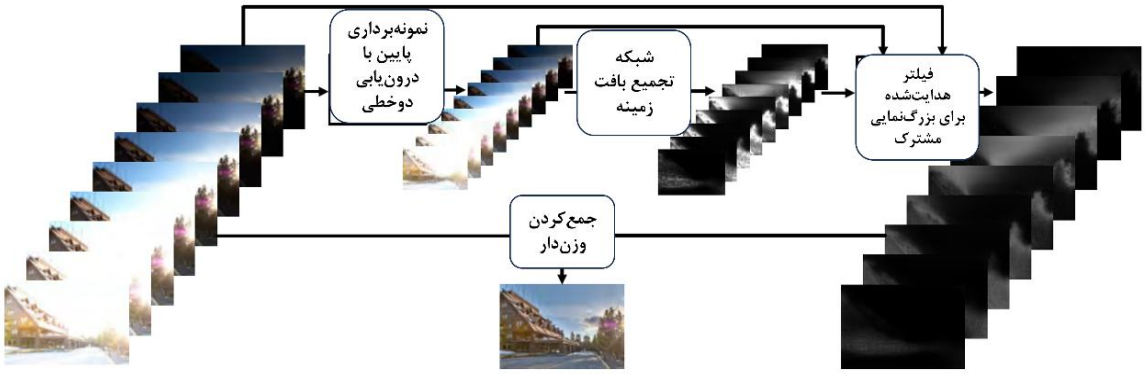
[5] *guided filter*

[Downloaded from jgit.kntu.ac.ir on 2026-01-29]

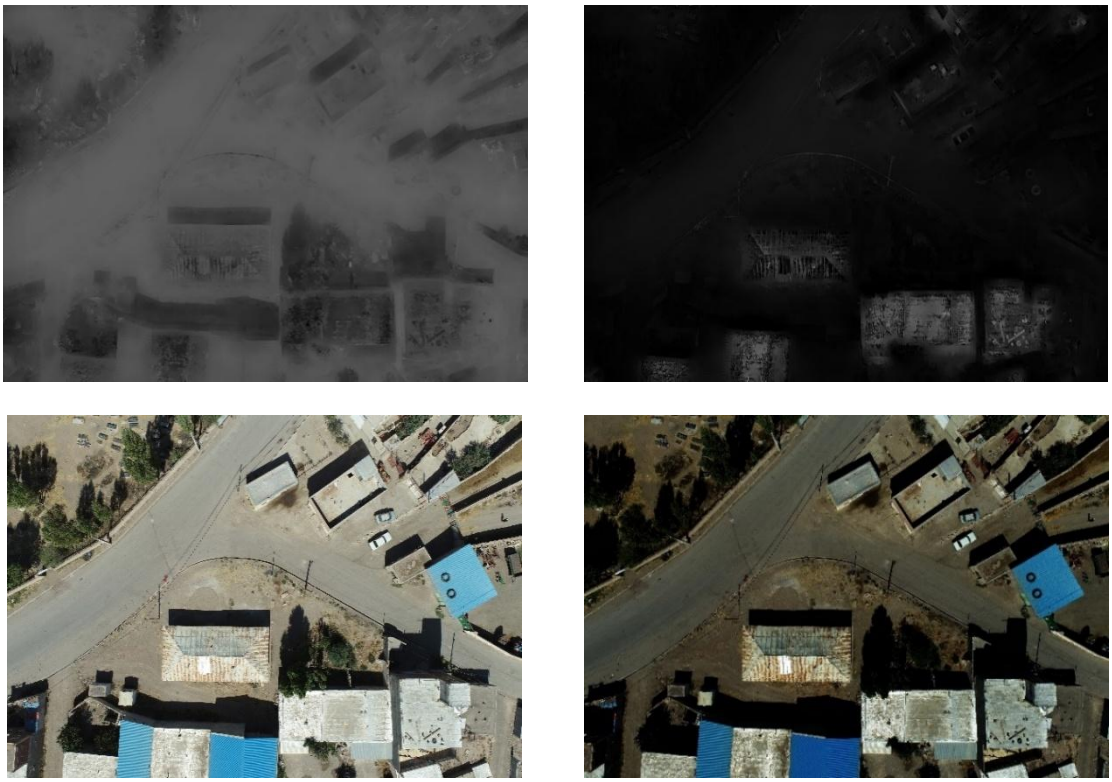
[DOI: 10.61882/jgit.13.1.77]

LeakyRelu در داخل شبکه استفاده می‌شود. این شبکه به‌منظور آموزش وزن‌های نقشه با رزولوشن پایین از ۷ لایه کانولوشن استفاده کرده است.

با توجه به شکل(۶) در نقشه‌های وزن هر چه تصویر روشن‌تر باشد وزن مقادیر روشن‌تر بالاتر است. به بیانی دیگر تأثیر مقادیر روشن‌تر نسبت به مقادیر تاریک‌تر بالاتر است. همان‌طور که در شکل(۵) ملاحظه می‌شود برای تصویر تاریک‌تر ناحیه‌هایی که بیش‌ازاندازه تاریک بوده دارای وزن پایین (تاریک‌تر) و برای تصویر روشن‌تر مقادیری که روشن‌تر است دارای وزن پایین (تاریک‌تر) می‌باشد که نشان می‌دهد در تصویر نهایی اثر مناطقی که دارای روشنایی بالا در تصویر روشن‌تر هستند کم‌تر بوده درحالی‌که این مناطق دارای وزن بالاتری در تصاویر تاریک دارند.



شکل۵: ساختار شبکه *MEF_NET*

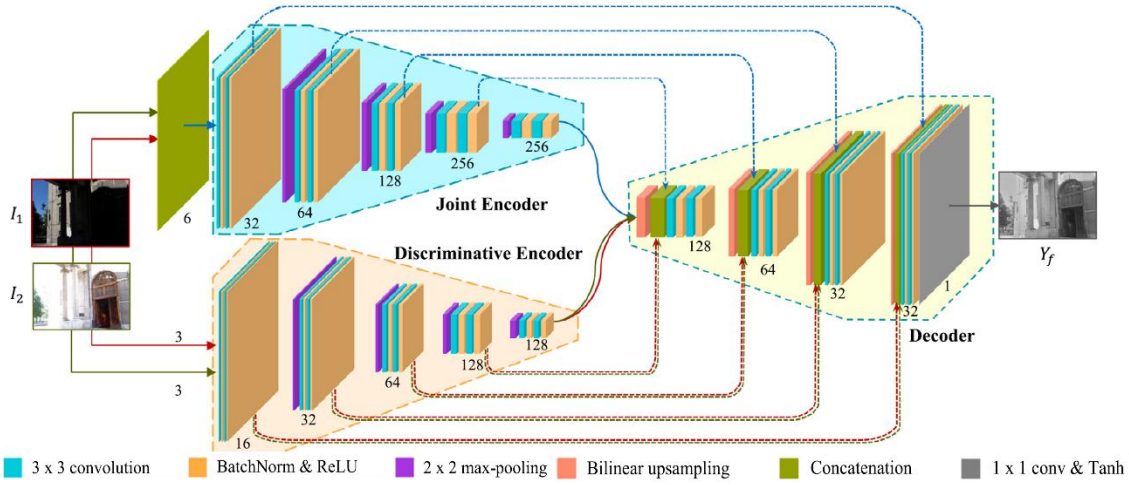


شکل۶: سطر اول: نمونه نقشه‌های وزن _ سطر دوم: نمونه تصویر اصلی به‌عنوان ورودی

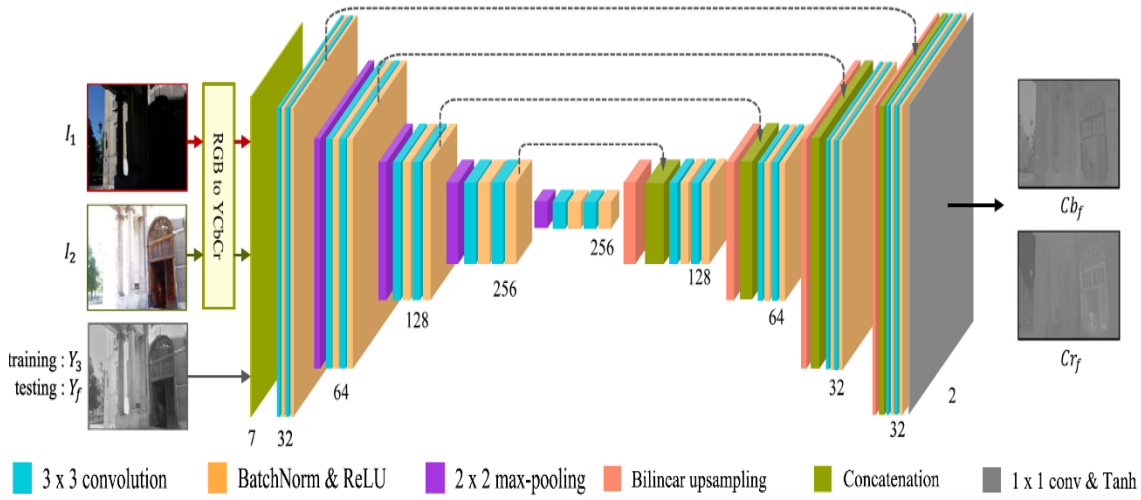
[Downloaded from jgit.kntu.ac.ir on 2026-01-29]
[DOI: 10.61882/jgit.13.1.77]

۴-۲- الگوریتم *DPE*

برخلاف الگوریتم *MEF*، الگوریتم *DPE* تنها دو ورودی می‌گیرد که شامل تصویر با روشنایی کم (*Under Exposure*) و روشنایی زیاد (*Over Exposure*) است. این روش از دو بخش کلی بهبود جزییات (*Detail enhancement module*) و بهبود رنگ (*Color enhancement module*) استفاده می‌کند (شکل‌های(۷) و (۸)). ابتدا در داخل خود شبکه در قسمت *DEM* (شکل(۷)) تصاویر جدیدی از تصاویر اصلی با روشنایی متفاوت تولید می‌شود. به‌منظور آموزش شبکه مربوط به *DEM* ابتدا از تصاویر اصلی، تصاویر شامل بیش‌ترین جزییات را با کمک ایجاد یک ماتریس نقشه مربوط به هر پیکسل و با الهام از نظریه رتینکس[۲۰] تولید می‌کند. ساختار شبکه *DEM* دارای دو بخش جداگانه رمزگذار_رمزگشا است که در اولی به دنبال همبستگی بالای ویژگی‌های بین دو تصویر و در بخش دیگر ویژگی‌های مربوط به هرکدام از تصاویر را جداگانه آموزش می‌دهد. هر تصویر به‌صورت جداگانه در ۵ لایه کانولوشن قرار می‌گیرد و سپس پس از کنار هم قرار دادن ویژگی‌ها در ۴ لایه کانولوشنی به سایز اصلی تصاویر برمی‌گردد و یک خروجی تصویر که دارای جزییات بالا است را نمایش می‌دهد.



شکل۷: ساختار شبکه بهبود جزییات

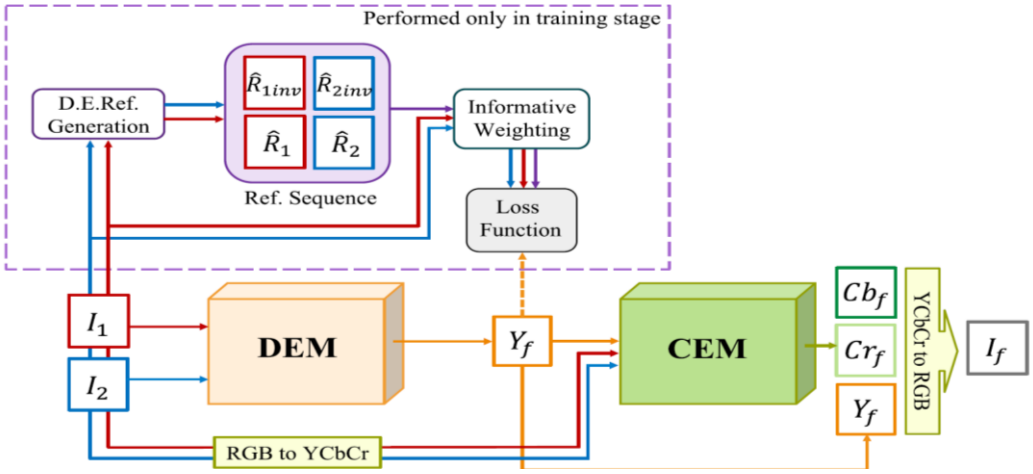


شکل۸: ساختار شبکه بهبود رنگ

[Downloaded from jgit.kntu.ac.ir on 2026-01-29]
[DOI: 10.61882/jgit.13.1.77]

برای شبکه *CEM* که هدف آن انتخاب مناسب‌ترین مقادیر کرومینانس تصویر است بار دیگر به دلیل عدم وجود داده‌های آموزشی مناسب از تصاویر موجود استفاده می‌کند. شبکه *CEM* در ۹ لایه کانولوشنی دو پارامتر مربوط به فضای رنگی *YCbCr* یعنی *Cb*و *Cr* را محاسبه می‌کند. در قسمت ورودی، دو تصویر رنگی و یک تصویر بهبودیافته تماما از یک صحنه از مرحله *DEM* در باند لومینانس وارد می‌شود. سپس برای تابع هزینه، از مقدار مرجع کرومینانس تصویر سوم که تنها در باند لومینانس وارد شده استفاده می‌شود و مدل آموزش می‌بیند. ساختار کلی این شبکه در شکل(۹) نشان داده شده است.

درنهایت از خروجی شبکه *CEM* مقادیر کرومینانس و از شبکه *DEM* مقادیر لومینانس به‌دست‌آمده و پس از تبدیل به باند رنگی *RGB* تصویر بهبودیافته تولید می‌شود.



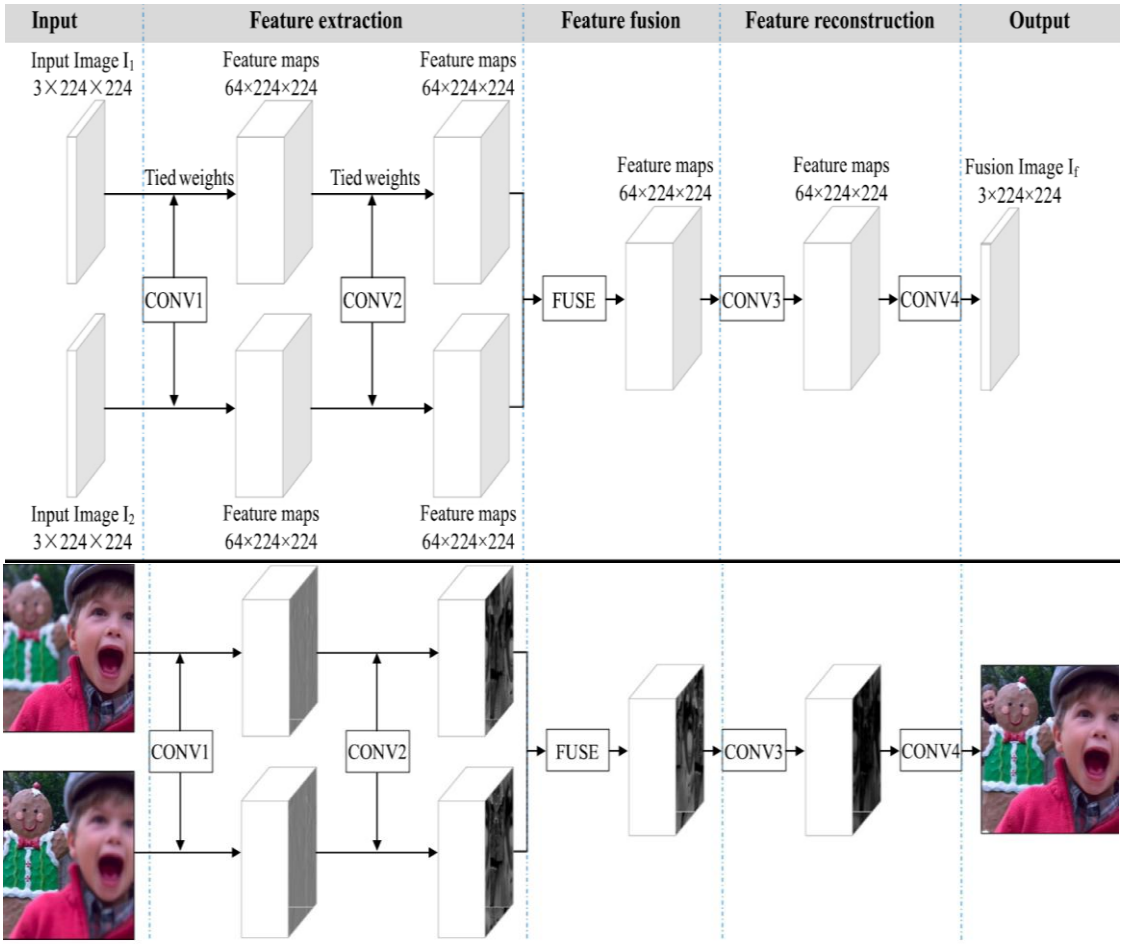
شکل۹: ساختار شبکه *DPE*

۴-۳- شبکه *IFCNN*

یکی دیگر از الگوریتم‌های توسعه‌یافته در حوزه بهبود روشنایی تصویر، *IFCNN* است. این شبکه یادگیری عمیق کانولوشنی می‌تواند علاوه بر بهبود روشنایی تصاویر، برای تصاویری همچون تصاویر چند فوکوس و مادون‌قرمز نیز مورداستفاده قرار گیرد. آموزش شبکه به کمک بیش از ۱۰۰ هزار تصویر چند فوکوس و یک مجموعه داده دارای تصاویر چند نوردهی انجام می‌شود. برای داده چند نوردهی تصاویر موجود از تعداد کمی برخوردار بوده که به بخش‌های کوچک‌تری تقسیم می‌شوند و شبکه به‌صورت نظارت‌نشده بر اساس شاخص *SSIM* آموزش‌دیده شده است. در دولایه اول، بدون کاهش اندازه تصویر، ویژگی‌های شاخص هر تصویر به‌صورت جدا استخراج می‌شود. سپس، با استفاده از روش قوانین ادغام عاقلانه عناصر[1] با یکدیگر ترکیب می‌شوند. پارامترهای کانولوشنی لایه اول بر اساس لایه کانولوشنی آموزش‌دیده شده از داده *ImageNet* ثابت در نظر گرفته‌شده است. در دولایه بعدی ویژگی‌ها بعد از ادغام تصاویر میزان شده[2] و درنهایت خروجی شبکه تولید می‌شود (شکل(۱۰)). برای تابع هدر رفت از دو پارامتر اختلاف تصویر تولیدشده با تصویر واقعی به همراه اختلاف نقشه ویژگی دو تصویر استفاده شده است.

[1] *element wise fusion rules*
[2] *tune*

[Downloaded from jgit.kntu.ac.ir on 2026-01-29]
[DOI: 10.61882/jgit.13.1.77]



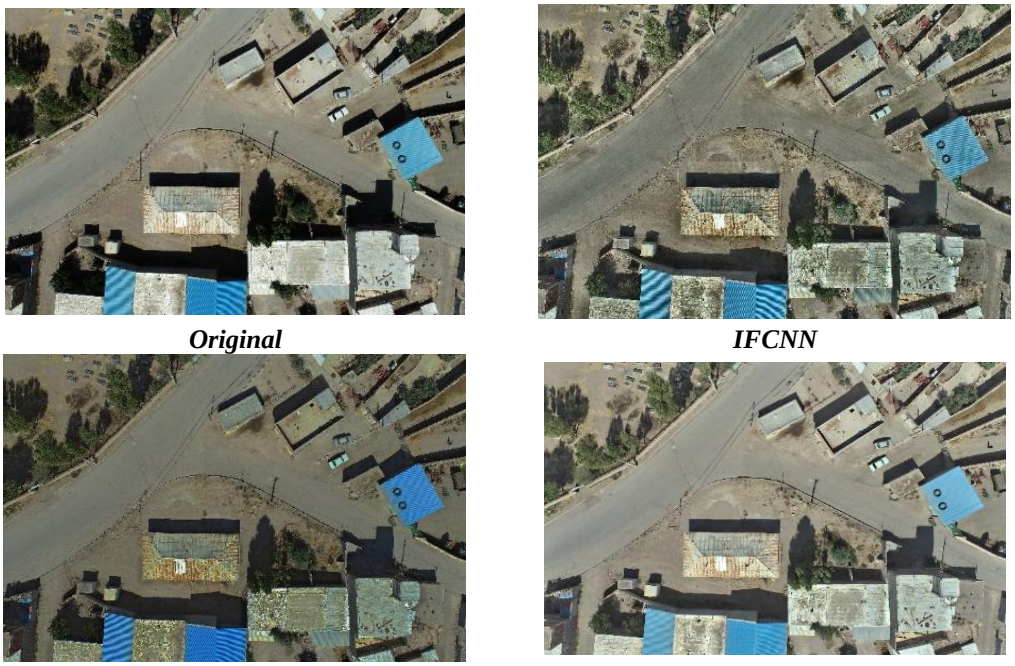
شکل۱۰: ساختار شبکه *IFCNN*

به دلیل رزولوشن بالای تصاویر هوایی نسبت به تصاویر مورداستفاده در شبکه‌های عصبی که معمولا از رزولوشن بسیار کمی استفاده می‌کنند از دو راه‌کار به‌منظور امکان بهره‌وری از شبکه‌های عصبی برای تصاویر با رزولوشن بالا استفاده شد. راه‌کار اول استفاده از سخت‌افزار قدرتمند با مقدار حافظه رم بالا و راه‌کار دوم تقسیم کردن تصویر به بخش‌های کوچک‌تر است. الگوریتم *MEF_NET* بر روی یک سیستم با مقدار حافظه ۳۲ گیگابایت و دو الگوریتم دیگر با تقسیم‌بندی تصاویر ورودی به تصاویر کوچک‌تر بر روی محیط کولب اجرا می‌شوند. به دلیل محدودیت مقدار حافظه در کولب، تصاویر ورودی به ۶ تصویر تقسیم‌بندی و سپس به‌عنوان ورودی به الگوریتم وارد می‌شوند. درنهایت خروجی‌های تولیدشده کنار یکدیگر قرارگرفته و تصویر نهایی تولید می‌گردد.

۵- نتایج تجربی

برای تصاویر منطقه مطالعاتی اول به‌منظور بهبود روشنایی در مناطق سایه نتایج به‌دست‌آمده از خروجی تصاویر هوایی بر روی شبکه‌های عصبی در شکل‌های(۱۱) و (۱۲) ملاحظه می‌شود.

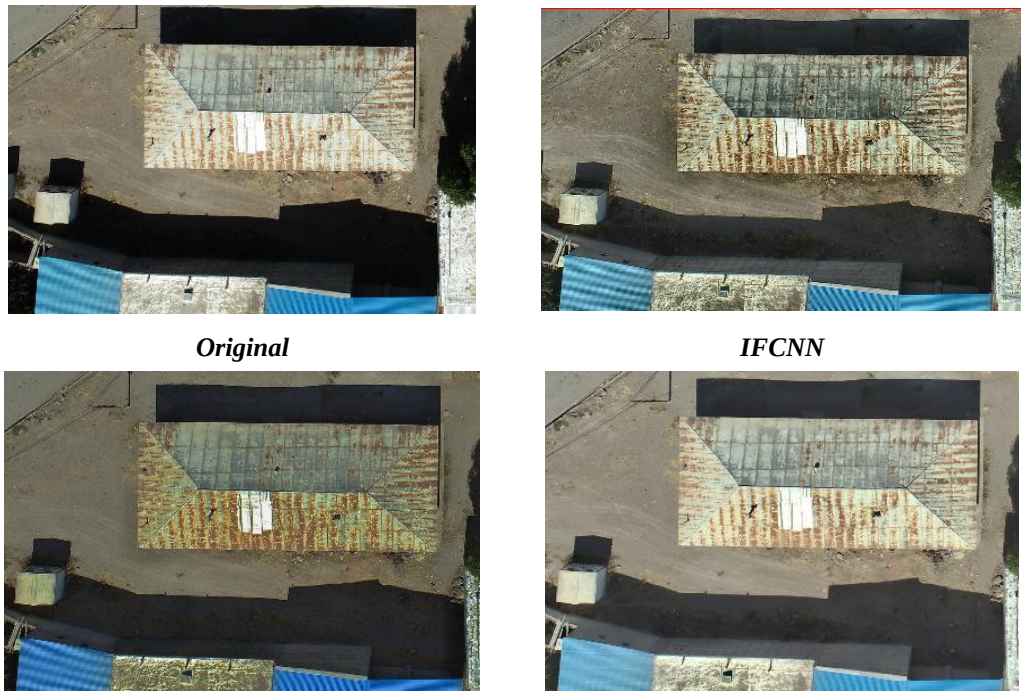
[Downloaded from jgit.kntu.ac.ir on 2026-01-29]
[DOI: 10.61882/jgit.13.1.77]



Original *IFCNN*

DEP *MEF*

شکل۱۱: نتایج خروجی برای تصاویر منطقه مطالعاتی اول



Original *IFCNN*

DEP *MEF*

شکل۱۲: بزرگنمایی نتایج برای تصاویر منطقه مطالعاتی اول

[Downloaded from jgit.kntu.ac.ir on 2026-01-29]

[DOI: 10.61882/jgit.13.1.77]

برای تصاویر منطقه مطالعاتی دوم، به‌منظور بازیابی اطلاعات در مناطق با روشنایی بالا، نتایج به‌دست‌آمده در شکل‌های(۱۳) و (۱۴) نشان داده شده است.

Original

IFCNN

DEP

MEF

شکل۱۳: نتایج خروجی برای تصاویر منطقه مطالعاتی دوم

Original

IFCNN

DEP

MEF

شکل۱۴: بزرگنمایی نتایج برای تصاویر منطقه مطالعاتی دوم

[Downloaded from jgit.kntu.ac.ir on 2026-01-29]

[DOI: 10.61882/jgit.13.1.77]

۶- ارزیابی نتایج

به‌منظور بررسی کیفیت تصویر بهبودیافته در تکنیک‌های مختلف بیان‌شده در قسمت‌های قبل، از ارزیابی معیارهای مبتنی بر تئوری اطلاعات یکپارچه شامل آنتروپی *(EN)* [۲۱]، سیگنال اوج نسبت به نویز *(PSNR)* [۲۲]، معیارهای یکپارچه مبتنی بر ویژگی تصویر شامل گرادیان متوسط *(AG)* [۲۳]، تطابق فاز (Q_P) [۲۴]، فرکانس فضایی *(SF)* [۲۵] معیارهای مبتنی بر شباهت شامل متریک (Q_C)[۲۶]، متریک یانگ (Q_Y) [۲۷]، اندازه‌گیری شاخص تشابه ساختاری *(MEF-SSIM)* *MEF* [۲۸]، معیارهای ادغام الهام گرفته از ادراک انسانی از ادراک بصری انسان شامل Q_{CB} [۲۹]، Q_{CV}[۳۰] استفاده می‌شود. نتایج بررسی تصاویر خروجی با شاخص‌های توضیح داده‌شده در جدول(۲) و جدول(۳) نشان داده شده است. نتایج شاخص‌ها بر اساس مقایسه بین تصویر بهبودیافته و تصویر مرجع می‌باشند. هرچه مقدار شاخص *Qcv* کمتر باشد، نشان‌دهنده عملکرد بهتر و در مابقی شاخص‌ها هرچه مقادیر بیشتر باشند، نشان از بهبود کارایی دارند. اعداد با رنگ قرمز بهترین نتیجه برای هر شاخص را نشان می‌دهند.

جدول۲: نتایج شاخص‌های به‌دست‌آمده برای تصویر دارای کمبود میزان روشنایی

شاخص / روش	Q_P	*AG*	*S_F*	Q_Y	*Qcv*	Q_{CB}	Q_C	*SSIM*	*EN*	*psnr*
MEF	۰/۴۸۹۳۴	۳/۸۴۰	۱۰/۰۹۱	۰/۶۹۷۴	۶۰/۲۷۸	۰/۲۵۸۶	۰/۶۹۷۴	۰/۹۲۸۶	۷/۲۹۰۳۰۲	۶۷/۰۸۴
DEP	۰/۴۷۳۳۸۲	۷/۱۹۹	۲۰/۰۵۶	۰/۴۶۹۹	۶۶/۳۱۳	۰/۲۴۳۸	۰/۴۶۹۹	۰/۸۸۸۲	۷/۰۳۴۲۱۸	۶۴/۷۵۹
IFCnn	۰/۰۱۰۱۴۳	۷/۱۹۹	۲۰/۰۵۶	۰/۰۴۱۰	۱۹۳/۵۲۱	۰/۱۷۸۲	۰/۰۱۶۴	۰/۲۵۲۱	۷/۰۳۳۲۱۸	۶۳/۲۹۱

جدول۳: نتایج شاخص‌های به‌دست‌آمده برای تصویر دارای روشنایی بیش‌ازاندازه

شاخص / روش	Q_P	*AG*	*S_F*	Q_Y	*Qcv*	Q_{CB}	Q_C	*SSIM*	*EN*	*psnr*
MEF	۰/۵۳۷۰	۳/۴۸۷۲	۸/۲۹۹۹	۰/۵۵۰۹	۲۱۹/۷	۰/۴۸۰۴	۰/۵۰۶۹	۰/۹۶۲۷	۶/۸۲۶۴۵۸	۶۶/۳۰۸
DEP	۰/۶۵۷۵	۶/۲۸۹۴	۱۷/۱۵۹	۰/۶۴۷۴	۲۱۸/۹	۰/۴۳۴۹	۰/۶۰۰۳	۰/۹۶۳	۷/۳۹۷۰۱	۶۶/۱۰۹
IFCnn	۰/۶۴۶۹	۶/۲۹۱۹	۱۶/۰۵۲	۰/۶۶۷۴	۳۷۱/۴	۰/۴۵۹۱	۰/۶۲۵۷	۰/۹۳۰	۷/۱۰۲۲۵	۶۶/۱۵۷

از بین تمام شاخص‌های مورد استفاده، رایج‌ترین شاخص، شاخص *SSIM*[۱۷] است. با توجه به نتایج منطقه مورد مطالعاتی اول (کمبود میزان روشنایی) در جدول(۲) روش *MEF_NET*، بهتر از دو روش دیگر عمل می‌کند. در منطقه مورد مطالعاتی دوم (روشنایی بیش‌ازاندازه) در جدول(۳) با اختلاف کمی در شاخص *SSIM* روش *DEP*، بهتر از دو روش دیگر عمل می‌کند.

در بررسی میزان زمان پردازش برای تصاویر، هر سه روش در زمان تقریبی ۱ دقیقه برای هر عکس با حجم ۸ مگابایتی(تعداد ۲۰ مگاپیکسلی) به طول انجامید. سخت‌افزار مورد استفاده برای الگوریتم *MEF_NET* دارای گرافیک *RTX*۳۰۵۰ با حافظه ۸ گیگ، رم ۳۲ گیگابایت و سی پی یو *k*۱۳۶۰۰ می‌باشد. دو الگوریتم دیگر بر روی محیط کولب اجرا شدند.

[Downloaded from jgit.kntu.ac.ir on 2026-01-29]
[DOI: 10.61882/jgit.13.1.77]

۷- نتیجه‌گیری

با توجه به الگوریتم‌های مورد ارزیابی به ترتیب، روش *DPE* با ۱۱ لایه پنهان، *MEF_NET* با ۷ لایه پنهان و *IFCNN* با ۴ لایه پنهان بیشترین پیچیدگی را داشتند. بر اساس نتایج به‌دست‌آمده توسط شاخص‌های روشنایی می‌توان بیان کرد که افزایش لایه‌های پنهان نتیجه بهتری را به همراه داشته است. الگوریتم *MEF_Net* با کمک استفاده از فیلترهای راهنما و تولید ماتریس‌های وزن برای تصاویر با میزان روشنایی متفاوت و انتخاب بهترین وزن برای هر پیکسل در تصویر ادغام‌شده، مناسب‌ترین میزان روشنایی برای هر پیکسل در تصویر نهایی را انتخاب می‌کند. در کنار این روش در الگوریتم *DPE* راهکار خود را به دو قسمت مختلف تقسیم می‌کند. در قسمت اول، ابتدا تصویر ادغام‌شده از ویژگی‌های تصاویر ورودی به‌دست‌آمده و سپس در قسمت دوم، همراه با تصاویر اصلی، روشنایی و رنگ آن بهبود می‌یابد. برخلاف سایر روش‌های دیگر که بهبود رنگ و روشنایی را جدا از انتخاب ویژگی تصاویر ورودی در الگوریتم خود در نظر نمی‌گیرند، در این روش در یک مرحله بیش‌تر و آموزش یک شبکه جدا، بهینه‌سازی روشنایی تصویر موردتوجه واقع‌شده است. شبکه *IFCNN* علی‌رغم عدم پیچیدگی در الگوریتم مورداستفاده و به کمک داده‌های بیش‌تر در یادگیری شبکه و بدون اعمال کاهش اندازه تصاویر در لایه‌های خود توانسته است نتایج مناسبی را در بهبود روشنایی تصاویر هوایی ارائه دهد.

در بین الگوریتم‌هایی که برای بهبود روشنایی تصویر با استفاده از یادگیری عمیق تاکنون منتشرشده است، ارزیابی دقیقی از میزان تأثیر این الگوریتم‌ها در تصاویر هوایی در شرایط مختلف نوری نشده است. با توجه به بررسی‌های انجام‌شده، ارزیابی بهبود روشنایی در تصاویر هوایی با استفاده از روش‌های سنتی تنها توسط زاور و شاه حسینی، مورد ارزیابی قرارگرفته است[۳۱]. از طرفی اکثر الگوریتم‌های منتشرشده در حوزه یادگیری عمیق روی‌داده‌های با تعداد پیکسل کم بهینه‌شده‌اند. به همین منظور، برای ارزیابی تصاویر با اندازه بالاتر از دو راهکار استفاده می‌شود. اولین راه‌کار استفاده از سخت‌افزار با میزان رم بالا و دومین راه‌کار تقسیم‌بندی، به تصاویر کوچک‌تر و سپس ادغام تصاویر نهایی در کنار هم و تولید تصویر اصلی پس از بهبود میزان روشنایی است. در این تحقیق، سعی شده است تا میزان اثربخشی سه الگوریتم یادگیری عمیق بر روی تصاویر هوایی در دو شرایط نوری کاملاً متفاوت، با استفاده از ۱۰ شاخص رایج در حوزه تصاویر با چند درجه روشنایی، ارزیابی شود. همچنین، میزان تفکیک مکانی یا رزولوشن تصاویر هوایی که از تعداد پیکسل بالایی برخوردار هستند، حفظ‌شده و با همان کیفیت تولید و بررسی می‌شود. در این تحقیق صرفا عملکرد شبکه‌های آموزش‌دیده شده مورد ارزیابی قرار گرفت اما می‌توان این شبکه‌ها را با استفاده از تصاویر هوایی آموزش داده و سپس عملکرد آن‌ها را موردبررسی قرارداد.

مراجع

[1] *X. Zhang, "Benchmarking and comparing multi-exposure image fusion algorithms," Information Fusion, vol. 74, pp. 111–131, 2021, doi: 10.1016/j.inffus.2021.02.005.*

[2] *Z. Ying, G. Li, and W. Gao, "A bio-inspired multi-exposure fusion framework for low-light image enhancement," arXiv preprint, arXiv:1711.00591, 2017. [Online]. Available: http://arxiv.org/abs/1711.00591*

[3] *K. Ma and Z. Wang, "Multi-exposure image fusion: A patch-wise approach," in Proc. IEEE Int. Conf. Image Process. (ICIP), 2015, pp. 1717–1721, doi: 10.1109/ICIP.2015.7351094.*

[4] *P. J. Burt and R. J. Kolczynski, "Enhanced image capture through fusion," in Proc. 4th Int. Conf. Computer Vision (ICCV), 1993, pp. 173–182, doi: 10.1109/ICCV.1993.378222.*

[5] *A. Vyas, S. Yu, and J. Paik, Fundamentals*

[Downloaded from jgit.kntu.ac.ir on 2026-01-29]

[DOI: 10.61882/jgit.13.1.77]

of Digital Image Processing, Signals and Communication Technology Series, pp. 3–11, 2018, doi: 10.1007/978-981-10-7272-7_1.

[6] A. M. Reza, "Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement," J. VLSI Signal Process. Syst. Signal Image Video Technol., vol. 38, no. 1, pp. 35–44, 2004, doi: 10.1023/B:VLSI.0000028532.53893.82.

[7] H. Ibrahim and N. S. P. Kong, "Brightness preserving dynamic histogram equalization for image contrast enhancement," IEEE Trans. Consum. Electron., vol. 53, no. 4, pp. 1752–1758, 2007, doi: 10.1109/TCE.2007.4429280.

[8] C. Wang and Z. Ye, "Brightness preserving histogram equalization with maximum entropy: A variational perspective," IEEE Trans. Consum. Electron., vol. 51, no. 4, pp. 1326–1334, 2005, doi: 10.1109/TCE.2005.1561863.

[9] N. Hayat and M. Imran, "Ghost-free multi exposure image fusion technique using dense SIFT descriptor and guided filter," J. Vis. Commun. Image Represent., vol. 62, pp. 295–308, 2019, doi: 10.1016/j.jvcir.2019.06.002.

[10] S. H. Lee, J. S. Park, and N. I. Cho, "A multi-exposure image fusion based on the adaptive weights reflecting the relative pixel intensity and global gradient," in Proc. IEEE Int. Conf. Image Process. (ICIP), 2018, pp. 1737–1741, doi: 10.1109/ICIP.2018.8451153.

[11] K. R. Prabhakar, V. S. Srikar, and R. V. Babu, "DeepFuse: A deep unsupervised approach for exposure fusion with extreme exposure image pairs," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), 2017, pp. 4724–4732, doi: 10.1109/ICCV.2017.505.

[12] H. Li and L. Zhang, "Multi-exposure fusion with CNN features," in Proc. IEEE Int. Conf. Image Process. (ICIP), 2018, pp. 1723–1727.

[13] J. Yin, B. Chen, Y. Peng, and C. Tsai, "Deep prior guided network for high-quality image fusion," 2020. [Online]. Available: https://arxiv.org/abs/2001.08941

[14] S. Y. Chen and Y. Y. Chuang, "Deep exposure fusion with deghosting via homography estimation and attention learning," in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), 2020, pp. 1464–1468, doi: 10.1109/ICASSP40776.2020.9053765.

[15] Z. Yang, Y. Chen, Z. Le, and Y. Ma, "GANFuse: A novel multi-exposure image fusion method based on generative adversarial networks," Neural Comput. Appl., vol. 33, no. 11, pp. 6133–6145, 2021, doi: 10.1007/s00521-020-05387-4.

[16] H. Xu, H. Liang, and J. Ma, "Unsupervised multi-exposure image fusion breaking exposure limits via contrastive learning," 2023.

[17] D. Han, L. Li, X. Guo, and J. Ma, "Multi-exposure image fusion via deep perceptual enhancement," Information Fusion, vol. 79, pp. 248–262, 2022, doi: 10.1016/j.inffus.2021.10.006.

[18] K. Ma, Z. Duanmu, H. Zhu, Y. Fang, and Z. Wang, "Deep guided learning for fast multi-exposure image fusion," IEEE Trans. Image Process., vol. 29, pp. 2808–2819, 2020, doi: 10.1109/TIP.2019.2952716.

[19] Y. Zhang et al., "IFCNN: A general image fusion framework based on convolutional neural network," Information Fusion, vol. 54, pp. 99–118, 2020, doi: 10.1016/j.inffus.2019.07.011.

[20] E. H. Land, "The Retinex theory of color vision," Scientific American, vol. 237, no. 6, pp. 108–128, 1977. [Online]. Available: https://www.semanticscholar.org/paper/The-Retinex-Theory-of-Color-Vision-SCIENTIFIC-Land/2f3f8f151a52afa3c1e80505ddb09b8624162e35

[21] J. W. Roberts, J. Van Aardt, and F. Ahmed, "Assessment of image fusion procedures

[Downloaded from jgit.kntu.ac.ir on 2026-01-29]

[DOI: 10.61882/jgit.13.1.77]

using entropy, image quality, and multispectral classification," J. Electron. Imaging, vol. 17, no. 2, pp. 1–28, 2008, doi: 10.1117/1.2945910.

[22] P. Jagalingam and A. Vittal, "A review of quality metrics for fused image," Aquatic Procedia, vol. 4, pp. 133–142, 2015, doi: 10.1016/j.aqpro.2015.02.019.

[23] G. Cui, H. Feng, Z. Xu, Q. Li, and Y. Chen, "Detail preserved fusion of visible and infrared images using regional saliency extraction and multi-scale image decomposition," Opt. Commun., vol. 341, pp. 199–209, 2015, doi: 10.1016/j.optcom.2014.12.032.

[24] I. Journal et al., "Image fusion based on an absolute feature," Int. J. Comput. Inf. Technol., vol. 3, no. 6, pp. 1433–1447, 2007.

[25] A. M. Eskicioglu and P. S. Fisher, "Image quality measures and their performance," IEEE Trans. Commun., vol. 43, no. 12, pp. 2959–2965, 1995.

[26] S. Pistonesi, J. Martinez, S. Mar, and R. Vallejos, "Structural similarity metrics for quality image fusion assessment," Image Process. On Line, 2018, doi: 10.5201/ipol.2018.196.

[27] S. Li, R. Hong, and X. Wu, "A novel similarity based quality metric for image fusion," in Proc. 3rd Int. Conf. Machine Learning and Cybernetics, 2008, pp. 167–172.

[28] K. Ma, K. Zeng, and Z. Wang, "Perceptual quality assessment for multi-exposure image fusion," IEEE Trans. Image Process., vol. 24, no. 11, pp. 3345–3356, 2015.

[29] Y. Chen and R. S. Blum, "A new automated quality assessment algorithm for image fusion," Image Vis. Comput., vol. 27, no. 10, pp. 1421–1432, 2009, doi: 10.1016/j.imavis.2007.12.002.

[30] H. Chen and P. K. Varshney, "A human perception inspired quality metric for image fusion based on regional information," Information Fusion, vol. 8, pp. 193–207, 2007, doi: 10.1016/j.inffus.2005.10.001.

[31] H. Zavar and R. Shah-Hosseini, "Comparative evaluation of lighting improvement methods in aerial images," J. Geospatial Inf. Technol., vol. 11, no. 3, pp. 103–119, Dec. 2023, doi: 10.61186/jgit.11.3.103.

[Downloaded from jgit.kntu.ac.ir on 2026-01-29]

[DOI: 10.61882/jgit.13.1.77]

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

Journal of Geospatial Information Technology
Vol.13, No.1, Spring 2025

Research Paper

Evaluation of the Enhancement of the Aerial Image Brightness Using Deep Learning Methods

Hossein Zavar [1], Reza Shah-Hosseini [2*]

1- PhD Student, Department of Photogrammetry and Remote Sensing, School of Surveying Engineering and Spatial Information, Faculty of Engineering, University of Tehran
2- Associate Professor, Department of Photogrammetry and Remote Sensing, School of Surveying Engineering and Spatial Information, Faculty of Engineering, University of Tehran

Abstract

Image taking using unmanned aerial vehicles (UAVs) for monitoring and assessing the existing conditions is one of the most prevalent applications in surveying. Despite its significant advantages, this approach also faces some challenges. Improper camera settings during image capturing, adverse weather conditions, and changes in lighting are the primary factors that reduce the quality of the captured images. Generally, the proposed methods for brightness enhancement can be categorized into two groups: traditional methods, which rely on histograms, and modern methods, including neural networks and deep learning, which have increasingly attracted the attention of researchers. The objective of this study is to evaluate the performance of deep learning methods in enhancing the brightness of the aerial images. These images often exhibit inadequate quality due to reduced visual details and the consequent loss of spectral information. Such deficiencies negatively impact the quality of the spatial products, such as orthophotos and digital surface models. Improving brightness and recovering spectral information have a direct and significant influence on the quality of these spatial products. To this end, the study examines three different deep learning methods that have demonstrated superior performance in the previous researches for enhancing aerial image brightness. The optimal method is selected based on 10 brightness evaluation metrics. The evaluated data consists of the aerial images captured from two different regions, characterized by areas with significant visual detail reduction and loss of spectral information due to poor lighting conditions. The results reveal hidden features in shadowed regions and areas with excessive brightness and high environmental reflection, which are not easily discernible by the naked eye. This is achieved by recovering spectral information through increasing the contrast between the digital values of the pixels in these regions. The best-performing method achieves structural similarity index (SSIM) scores of 0.92 and 0.96 for the two datasets, respectively. SSIM is one of the most critical evaluation metrics among the 10 criteria utilized in this study.

Key words: *Photogrammetry, Fusion, Brightness Enhancement, Orthophoto, Deep Learning.*

Correspondence Address*: Department of Photogrammetry and Remote Sensing, School of Surveying Engineering and Spatial Information, Faculty of Engineering, University of Tehran,Tehran, Iran.*
Tel*: +98 21 61114527*
Email: *rshahosseini@ut.ac.ir*

[Downloaded from jgit.kntu.ac.ir on 2026-01-29]

[DOI: 10.61882/jgit.13.1.77]