



Research Paper

Shadow Correction in UAV Imagery Using a Deep Convolutional U-Net Neural Network

Fardin Fazeli Gargari¹, Ali Shams-al-Dini², Amir Aghabalai^{*3}, Ali Mohtadi⁴, Mehdi Haji Ghasemi⁵

1. Master of Science, Khajeh Nasir al-Din Tusi University of Technology
2. Associate Professor, Remote Sensing Department, Tarbiat Modares University
3. Assistant Professor, Imam Hussein (AS) University
4. Hekmat Institute of Higher Education, Qom
5. Faculty Member, University of Tehran

ABSTRACT

Shadows in unmanned aerial vehicle (UAV) imagery introduce significant visual and radiometric distortions that adversely affect feature extraction, classification, and three-dimensional reconstruction accuracy. To mitigate these effects, this study proposes a deep learning-based shadow correction framework built upon a customized U-Net architecture enhanced with a VGG19-based encoder. The model was trained on paired UAV images (shadowed and shadow-free) normalized to the $[0,1]$ intensity range. Network parameters were optimized using the ADAM optimizer and a hybrid loss function combining mean absolute error (MAE) and mean squared error (MSE) to balance pixel-level accuracy and structural preservation.

Quantitative evaluation using Root Mean Square Error (RMSE) and Peak Signal-to-Noise Ratio (PSNR) demonstrated that the proposed model achieved an RMSE of 0.0404 in the normalized $[0,1]$ range (equivalent to 10.31 in the 8-bit $[0-255]$ scale) and a PSNR of 27.87 dB, indicating accurate reconstruction of shadowed regions while preserving fine structural and textural details. Qualitative assessments further confirmed stable performance across high-resolution short-range imagery without introducing noticeable artifacts.

The results suggest that enhanced convolutional architectures remain highly effective for shadow correction in high-detail UAV scenarios, providing a favorable balance between reconstruction accuracy and computational efficiency. Integrating such deep convolutional frameworks into UAV image preprocessing pipelines can significantly improve radiometric consistency and the analytical reliability of photogrammetric datasets.

UAV imagery; shadow correction; deep learning; U-Net.

Keywords:

*Corresponding Author: Safarahmati@modares.ac.ir

ORCID ID: <https://orcid.org/0000-0001-5774-7154>



Copyright© 2025, TMU Press. This open-access article is published under the terms of the Creative Commons Attribution- NonCommercial 4.0 International License which permits Share (copy and redistribute the material in any medium or format) and Adapt (remix, transform, and build upon the material) under the Attribution-NonCommercial terms.

کاهش اثر سایه در تصاویر پهپاد با استفاده از شبکه عصبی عمیق U-Net

فردین فاضلی گرگری^۱، علی شمس الدینی^۲، امیر آقابالائی^{۳*}، علی مهتدی^۴،
مهتدی حاجی قاسمی^۵

۱. کارشناس ارشد، دانشگاه صنعتی خواجه نصیرالدین طوسی
۲. دانشیار گروه سنجش از دور، دانشگاه تربیت مدرس
۳. استادیار دانشگاه جامع امام حسین (ع)
۴. موسسه آموزش عالی حکمت قم
۵. عضو هیات علمی دانشگاه تهران

چکیده

وجود سایه در تصاویر پهپادی سبب بروز اعوجاج‌های بصری و پرتوی می‌شود که صحت استخراج ویژگی‌ها، طبقه‌بندی و بازسازی سه‌بعدی را به‌طور قابل‌توجهی کاهش می‌دهد. این پژوهش با هدف کاهش اثرات سایه، چارچوبی مبتنی بر یادگیری عمیق را برای تصحیح سایه در تصاویر پهپادی ارائه می‌کند که در آن از معماری شبکه عصبی کانولوشنی U-Net سفارشی‌سازی شده استفاده شده است. مدل پیشنهادی با ۳۲۰ جفت تصویر دارای سایه و بدون سایه آموزش داده شد و داده‌ها پیش از آموزش به بازه‌ی [۰،۱] نرمال‌سازی گردیدند و عملکرد مدل با استفاده از شاخص‌های ریشه میانگین مربعات خطا (RMSE)، نسبت بیشینه سیگنال به نویز (PSNR) و شاخص شباهت ساختاری (SSIM) بر روی مجموعه آزمون ارزیابی شد. به‌منظور بهینه‌سازی فرآیند یادگیری، از بهینه‌ساز ADAM و تابع خطای میانگین قدرمطلق (MAE) استفاده شده است. نتایج کمی حاصل از شاخص‌های RMSE و PSNR نشان داد که مدل پیشنهادی توانسته است به مقدار RMSE برابر با ۰/۰۴۰۴ در مقیاس نرمال‌شده [۰،۱] (معادل ۱۰/۳۱ در مقیاس ۸ بیتی [۰،۲۵۵]) و PSNR برابر با ۲۷/۸۷ دسی‌بل دست یابد که بیانگر بازسازی دقیق نواحی سایه‌دار همراه با حفظ ساختارهای بافتی و هندسی تصویر است. ارزیابی کیفی نیز نشان داد که مدل در تصاویر با جزئیات بالا عملکردی پایدار ارائه داده و از بروز آرتیفکت‌های محسوس جلوگیری می‌کند. به‌طور کلی، به‌کارگیری معماری‌های عمیق کانولوشنی در فرایند پیش‌پردازش تصاویر پهپادی می‌تواند یکنواختی پرتوی و قابلیت اعتماد تحلیلی داده‌های فتوگرامتری را به‌طور معناداری بهبود بخشد.

تصاویر پهپادی، تصحیح سایه، یادگیری عمیق، شبکه U-Net

واژگان کلیدی:

remotesensing@ihu.ac.ir

* نویسنده مسئول

کپی‌رایت © ۲۰۲۵، انتشارات دانشگاه تربیت مدرس (TMU Press). این مقاله به‌صورت دسترسی آزاد منتشر شده و تحت مجوز بین‌المللی Creative Commons Attribution-NonCommercial 4.0 قرار دارد. بر اساس این مجوز، شما می‌توانید این مطلب را در هر قالب و رسانه‌ای کپی، بازنشر و بازآفرینی کنید و یا آن را ویرایش و بازسازی نمایید، به شرط آنکه نام نویسنده را ذکر کرده و از آن برای مقاصد غیرتجاری استفاده کنید.

۱. مقدمه

استفاده از پهپادها^۱ در فتوگرامتری، نقشه‌برداری و پایش محیط‌زیست طی سال‌های اخیر با رشد چشمگیری همراه بوده است (Saadatseresht et al., 2015; Kern et al., 2021). پهپادها به دلیل هزینه‌ی پایین، سرعت بالای تصویربرداری و توانایی تولید داده‌های مکانی با قدرت تفکیک بالا، ابزاری کارآمد در تولید مدل‌های رقومی سطح و نقشه‌های دقیق محسوب می‌شوند (Motayyeb et al., 2022). با این وجود، داده‌های تصویری حاصل از پهپادها به شدت تحت تأثیر شرایط نوری، ارتفاع پرواز و زاویه‌ی تابش قرار دارند؛ عواملی که می‌توانند موجب بروز سایه‌های پیچیده^۲ در سطح زمین شوند (Khan et al., 2016; Zali et al., 2022).

پدیده‌ی سایه در تصاویر هوایی و پهپادی، از چالش‌های بنیادی در فتوگرامتری و سنجش‌ازدور است. وجود سایه‌ها که در اثر انسداد تابش نور توسط موانع طبیعی یا مصنوعی ایجاد می‌شود، موجب کاهش روشنایی، تغییر رنگ و از دست رفتن اطلاعات طیفی می‌گردد. این تغییرات در نهایت باعث بروز ناهمگنی پرتوی^۳ شده و صحت عملیات‌هایی چون استخراج ویژگی، بازسازی سه‌بعدی، و تفسیر عوارض زمینی را کاهش می‌دهد (Finlayson et al., 2006; Movia et al., 2016). این پدیده به‌ویژه در محیط‌های شهری و ناهموار، که تنوع ارتفاعی زیاد است، به شکل شدیدتر بروز می‌کند (Dare, 2005).

همچنین نشان داده شده است که وجود سایه علاوه بر کاهش دقت مکانی، در محاسبه شاخص‌های طیفی مانند NDVI و شاخص‌های پوشش گیاهی نیز انحراف ایجاد می‌کند (Saadatseresht et al., 2015; Alberto et al., 2016; Xu et al., 2019; Movia et al., 2016). از این رو، کاهش اثر سایه به‌عنوان یکی از مراحل کلیدی در پیش‌پردازش داده‌های UAV مطرح شده است (Zali et al., 2022).

در مطالعات جدیدتر، تأکید شده است که در برداشت‌های پهپادی، تغییرات روشنایی و اثرات سایه نه فقط کیفیت بصری، بلکه «سازگاری پرتوی بین تصاویر» در مجموعه‌های همپوشان را نیز مختل می‌کند و در نتیجه می‌تواند کیفیت اورتوموزاییک، شاخص‌های طیفی و محصولات SfM را به‌طور سیستماتیک تحت تأثیر قرار دهد. از این رو، راهکارهای نرمال‌سازی پرتوی و تنظیمات پرواز برای کاهش سایه، به‌عنوان بخش مکمل پیش‌پردازش و حتی طراحی مأموریت برداشت داده مطرح شده‌اند (Maes, 2025). همچنین پژوهش‌های ۲۰۲۴ نشان می‌دهند راهبردهای نرمال‌سازی پرتوی در تصاویر هوایی و پهپادی، به‌طور مشخص برای رفع ناسازگاری‌های روشنایی درون‌تصویری و بین‌تصویری توسعه یافته‌اند و می‌توانند خطاهای ناشی از تغییر شرایط نور و اثرات بازتابی را کاهش دهند (Liu et al., 2024). در کاربردهای کشاورزی دقیق نیز گزارش شده است که کالیبراسیون پرتوی مبتنی بر سنجش تابش ورودی، می‌تواند اثر سایه و تغییرات گذرای روشنایی را در بازتابندگی برآوردشده تعدیل کند (Swaminathan et al., 2024).

در نخستین تلاش‌ها برای مقابله با این پدیده، روش‌های کلاسیک پردازش تصویر بر مبنای ویژگی‌های رنگ و روشنایی طراحی شدند. الگوریتم‌های مبتنی بر مدل Retinex از شناخته‌شده‌ترین آن‌ها هستند که شدت روشنایی هر پیکسل را

1 Unmanned Aerial Vehicles; UAVs

2 Complex Shadows

3 Radiometric Inconsistency

حاصل ضرب بازتاب و روشنایی محیط در نظر می‌گیرند و با برآورد این مؤلفه‌ها، سعی در بازسازی تصویر بدون سایه دارند (Finlayson et al., 2006; Drew & Nr, 2005; Yang et al., 2023). در کنار آن، روش‌هایی مانند آستانه‌گذاری تطبیقی^۱ و متعادل سازی هیستوگرام^۲ به منظور یکنواخت‌سازی روشنایی در نواحی سایه‌دار به کار گرفته شدند. این روش‌ها اگرچه در صحنه‌های ساده و با سایه‌های واضح عملکرد قابل قبولی دارند، اما در محیط‌هایی با روشنایی ناهمگون، سایه‌های نیمه‌شفاف یا بازتاب‌های پیچیده کارایی خود را از دست می‌دهند و معمولاً باعث بیش‌روشن‌سازی یا از بین رفتن بافت تصویر می‌شوند (Dare, 2005; Movia et al., 2016).

در امتداد این بحث، پژوهش‌های ۲۰۲۴ در حوزه تصاویر هوایی با وضوح بالا نشان داده‌اند که حتی رویکردهای غیرعمیق یا نیمه‌کلاسیک نیز وقتی با منطق «همبستگی ناحیه‌ای» و بازبایی بافت مبتنی بر گروه‌بندی عوارض ترکیب شوند، می‌توانند در صحنه‌های پیچیده کارایی بهتری نسبت به همسان‌سازی‌های عمومی روشنایی داشته باشند و به‌ویژه در مرزهای سایه، پیوستگی طبیعی‌تری ایجاد کنند (Guo et al., 2024).

ظهور یادگیری عمیق^۳ نقطه‌ی عطفی در حل این مسئله بود. شبکه‌های عصبی کانولوشنی (CNN) توانستند به‌جای تکیه بر قواعد از پیش تعریف‌شده، روابط مکانی و پرتوی را به‌صورت داده‌محور یاد بگیرند (Kim, 2019; Lecun et al., 2015). در این میان، U-Net به‌عنوان یکی از معماری‌های مؤثر در بازسازی پیکسلی معرفی شد (Guo et al., 2020; Member, 2024). ساختار رمزگذار-رمزگشا و وجود اتصالات پرش در این شبکه سبب می‌شود جزئیات مکانی که در لایه‌های عمیق از بین می‌روند، به مراحل بازسازی منتقل شوند و بدین ترتیب، نواحی سایه‌دار بدون از دست دادن ساختار و بافت اصلی، بازسازی گردند (Hu et al., 2018). مطالعات مختلف، از جمله در حوزه‌ی تصحیح سایه در داده‌های UAV، کارایی بالای این معماری را تأیید کرده‌اند (Zali et al., 2022; Saadatseresht et al., 2015; Kern et al., 2021; Motayyeb et al., 2022).

به‌منظور ارتقای بیشتر عملکرد مدل، نسخه‌های پیشرفته‌تری از U-Net معرفی شده‌اند. برای نمونه، مدل‌های Attention U-Net با افزودن سازوکار تمرکز^۴ امکان تشخیص دقیق‌تر نواحی سایه‌دار را فراهم کرده‌اند (Iakovidis & Member, 2024; Nagee et al., 2021). در رویکردی دیگر، ترکیب شبکه‌ی U-Net با شبکه‌های مولد تخصصی^۵ نظیر Mask-ShadowGAN^۶ و RIS-GAN^۶ باعث شده است بازسازی مرزهای سایه و مناطق نیمه‌شفاف با صحت بالاتری انجام شود (Hu et al., 2019; Zhang et al., 2020). این مدل‌ها با یادگیری توزیع آماری روشنایی در تصاویر واقعی، قادرند تفاوت میان نواحی سایه‌دار و روشن را به‌صورت خودکار تشخیص دهند و بافت‌های از دست‌رفته را بازتولید کنند، هرچند گاهی به پایداری آموزشی حساس هستند (Nagee et al., 2021; Ding et al., 2019).

1 Adaptive Thresholding

2 Histogram Equalization

3 Deep Learning

4 Attention Mechanism

5 GANs

6 Region Interaction Shadow Generative Adversarial Network

با گسترش کاربرد ترنسفورمرها در بینایی ماشین، پژوهش‌ها به سمت مدل‌هایی رفتند که بتوانند وابستگی‌های بلندبرد^۱ در تصویر را مدل کنند. TSRFormer نمونه‌ای از این رویکردهاست که با طراحی یک چارچوب دومرحله‌ای مبتنی بر توجه سراسری، ابتدا نواحی سایه را آشکار و سپس بازسازی پرتوی را انجام می‌دهد (Chang et al., 2023). چنین مدل‌هایی به دلیل دید سراسری خود، توانایی بهتری در سازگاری روشنایی بین بخش‌های سایه‌دار و غیرسایه‌دار دارند، هرچند هزینه محاسباتی آن‌ها نسبت به شبکه‌های کانولوشنی بیشتر است (Zhang et al., 2018).

در سال ۲۰۲۵، مسیر دیگری نیز تقویت شده است که به جای اتکا به Transformer خالص، از طراحی‌های هیبریدی بهره می‌گیرد تا ضمن حفظ دید سراسری، هزینه محاسباتی کنترل شود. برای نمونه، یک مدل ترکیبی Transformer و GAN با سازوکار بازوزن‌دهی توجه، با هدف افزایش سازگاری روشنایی بین ناحیه سایه و غیرسایه و همزمان کاهش پیچیدگی، گزارش شده و بهبودهای معناداری را در معیارهای PSNR و RMSE روی مجموعه داده‌های مرجع حذف سایه نشان داده است (Hu et al., 2025). این خط پژوهش، با دغدغه کاربردپذیری در سامانه‌های محدود از نظر توان پردازشی همسو است و با ادبیات «پردازش سبک و نزدیک به بلادرنگ» در بینایی ماشین پهپادی هم‌راستا تلقی می‌شود (Piecznyński et al., 2024).

در سال‌های اخیر، مدل‌های انتشاری^۲ به‌عنوان نسل تازه‌ای از شبکه‌های مولد برای بازسازی تصویر معرفی شده‌اند که با فرایند تدریجی و تکرارشونده، قادر به تولید تصاویر با جزئیات دقیق و روشنایی یکنواخت هستند (Luo et al., 2024; Mei et al., 2023). این مدل‌ها در مقایسه با شبکه‌های مولد تخصصی پایداری آموزشی بالاتری داشته و در ترکیب با ساختارهایی مانند U-Net و ماژول‌های توجه سراسری، نتایج قابل قبولی در حوزه‌ی کاهش اثر سایه کسب کرده‌اند (Vasluianu et al., 2025). در چالش NTIRE 2025، مدل‌های ترکیبی مبتنی بر ترنسفورمر و انتشار توانستند به بالاترین مقادیر PSNR^۳ و SSIM^۴ گزارش شده دست یابند (Vasluianu et al., 2025). گزارش رسمی چالش NTIRE 2025 نشان می‌دهد تمرکز بسیاری از راهکارهای برتر بر ترکیب بازسازی مولد و سازگاری پرتوی است و ارزیابی‌ها در دو مسیر «وفاداری بازسازی» و «ادراک بصری» انجام شده است که به‌طور ضمنی اهمیت کنترل یکنواختی روشنایی و کاهش مرزهای مصنوعی در خروجی را برجسته می‌کند (Vasluianu et al., 2025). در همین راستا، مدل‌های مبتنی بر انتشار نیز با هدف کاهش ناسازگاری روشنایی و ارتقای کیفیت بازسازی معرفی شده‌اند و یک نمونه شاخص آن، رویکرد انتشار با راهبری سراسری است که برای کاهش مرزهای غیرطبیعی و بهبود یکنواختی روشنایی در حذف سایه پیشنهاد شده است (Luo et al., 2024). همزمان، به دلیل دشواری تولید داده‌های جفت‌شده در سناریوهای واقعی پهپادی، پژوهش‌های ۲۰۲۵ به سمت حذف سایه با داده‌های unpaired حرکت کرده‌اند و راهکارهایی برای تقویت توجه به نواحی سایه در چارچوب‌های بدون جفت ارائه شده است (Yang & Shin, 2025). توسعه‌ی معماری‌های سبک‌تر مانند SpA-Former^۵، با هدف کاهش پیچیدگی محاسباتی برای کاربردهای پهپادی، مسیر جدیدی برای استفاده‌ی عملی از این مدل‌ها گشوده است

1 Long-range Dependencies

2 Diffusion Models

3 Peak Signal-to-Noise Ratio

4 Structural Similarity Index Measure

5 Shadow-aware Position Attention Transformer

(Zhang, 2023).

در سال‌های اخیر، با پیشرفت‌های چشمگیر در یادگیری عمیق تمرکز پژوهش‌ها از مدل‌های مبتنی بر قوانین ثابت به سمت شبکه‌های عصبی کانولوشنی^۱ سوق یافته است (Kim, 2019; Lecun et al., 2015). این شبکه‌ها قادرند الگوهای غیرخطی پیچیده میان پیکسل‌ها را یاد بگیرند و از طریق لایه‌های متوالی کانولوشن و غیرخطی سازی، روابط مکانی و طیفی را بازسازی کنند. گو^۲ و همکاران (۲۰۲۰) با بهره‌گیری از U-Net همراه با بلوک توجه^۳ و چندین تابع خطا، استخراج ساختمان‌ها را بهبود دادند، در حالی که یاکووی دیس و همکاران (۲۰۲۴) مدل Soft-Hard Attention U-Net را برای حذف سایه‌های چندمقیاسی معرفی کردند. شبکه‌ی U-Net از دو مسیر متقارن تشکیل شده است: رمزگذار^۴ برای استخراج ویژگی‌های سطحی و عمیق، و رمزگشا^۵ برای بازسازی تدریجی تصویر.

اتصالات پرش^۶ میان این دو مسیر موجب می‌شود جزئیات مکانی از لایه‌های اولیه به مراحل بازسازی منتقل شوند، در نتیجه مدل می‌تواند نواحی سایه‌دار را بدون از بین بردن بافت و ساختار تصویر بازسازی کند (Hu et al., 2018). این ویژگی سبب شده است U-Net در کاربردهای مرتبط با بهبود تصاویر هوایی، از جمله حذف سایه، افزایش وضوح و استخراج عوارض، عملکرد مناسبی داشته باشد (Zali et al., 2022).

با وجود این پیشرفت‌ها، چند چالش اساسی همچنان باقی است. یکی از این چالش‌ها، وابستگی بسیاری از شبکه‌ها به داده‌های جفت‌شده‌ی دارای سایه و فاقد سایه^۷ است؛ داده‌ای که تهیه‌ی آن برای تصاویر پهپاد پرهزینه و دشوار است (Shen et al., 2017; Qu et al., 2017; Hua & Xia, 2019; Xu et al., 2019).

چالش دیگر، حساسیت زیاد عملکرد مدل‌ها به پارامترهای آموزشی^۸ همچون نرخ یادگیری و اندازه‌ی دسته^۹ است که انتخاب نادرست آن‌ها منجر به بیش‌برازش^{۱۰} یا همگرایی ناقص می‌شود (Zhang et al., 2018; Qu et al., 2017). چالش دیگر، ضرورت ایجاد تعادل میان صحت و سبکی مدل برای کاربردهای واقعی پهپادی است؛ چراکه سامانه‌های پهپادی اغلب از توان محاسباتی محدودی برخوردارند و به الگوریتم‌هایی نیاز دارند که با پیچیدگی کمتر، نتایج با کیفیت بالا ارائه دهند (Kern et al., 2021).

مرور ادبیات نشان می‌دهد مسیر تکامل علمی در این حوزه از مدل‌های فیزیکی ساده مانند Retinex و همسان‌سازی هیستوگرام (Finlayson et al., 2006; Drew & Nr, 2005; Yang et al., 2023) به سمت شبکه‌های یادگیری عمیق (U-Net, GAN و Transformer) و نهایتاً مدل‌های ترکیبی و مولد انتشاری حرکت کرده است (Zali et al., 2022; Guo et al., 2020; Chang et al., 2023; Nagae et al., 2021; Hu et al., 2019; Zhang et al., 2020; Luo et

1 CNNs

2 Guo

3 Attention Blocks

4 Encoder

5 Decoder

6 Skip Connections

7 Shadow/Shadow-free pairs

8 Hyperparameters

9 Batch Size

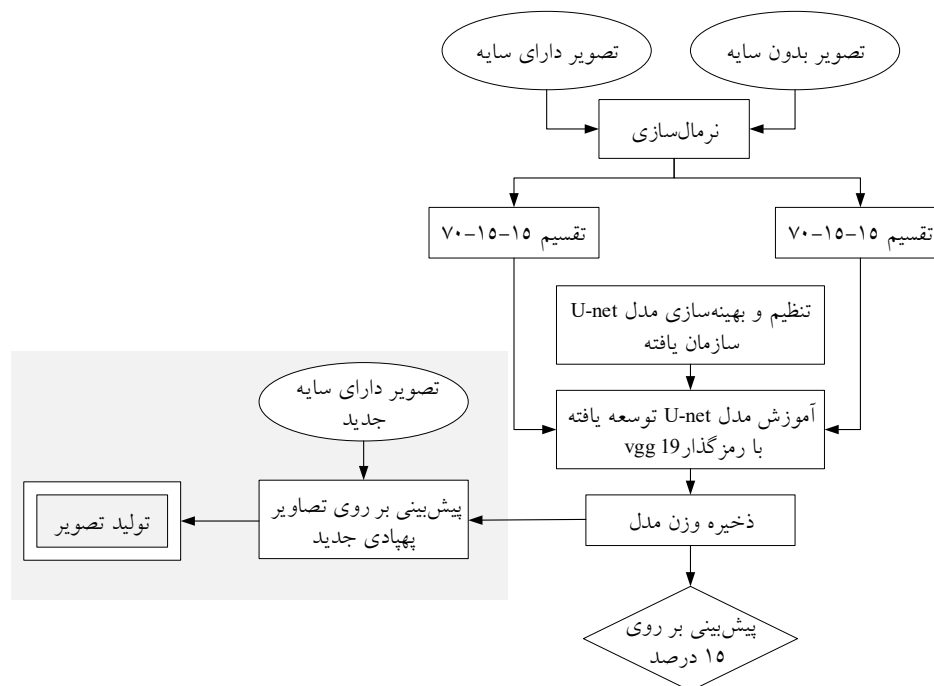
10 Overfitting

al., 2024–Vasluianu et al., 2025). پژوهش حاضر در ادامه‌ی این روند، تلاش دارد با بهره‌گیری از ظرفیت‌های U-Net و تحلیل اثر پارامترهای آموزشی، رویکردی پایدار و بهینه برای تعدیل اثر سایه در تصاویر پهپادی ارائه دهد که ضمن حفظ جزئیات مکانی، یکنواختی پرتوی را نیز بهبود بخشد.

۲. مواد و روش‌ها

فرایند کلی پژوهش در پنج مرحله اصلی سازمان‌دهی شده است:

مرحله اول شامل گردآوری داده‌های جفت‌شده دارای سایه و بدون سایه و بدون سایه از صحنه‌های واقعی با استفاده از پهپاد مجهز به دوربین RGB با وضوح بالا می‌باشد. سپس عملیات پیش‌پردازش شامل نرمال‌سازی شدت روشنایی، تفکیک داده‌ها به مجموعه‌های آموزش، اعتبارسنجی و آزمون (به نسبت ۷۰٪، ۱۵٪، ۱۵٪)، انجام گردید و روش‌های افزایش داده به‌منظور ارتقای تعمیم‌پذیری مدل اعمال گردید. در مرحله بعد، طراحی و پیاده‌سازی معماری بهبودیافته شبکه U-Net با استفاده از بلوک‌های انتقال و ویژگی مبتنی بر VGG19¹ و ماژول‌های توجه کانالی انجام شد. سپس، آموزش مدل با بهره‌گیری از الگوریتم Adam Optimizer، تابع هزینه میانگین قدرمطلق خطا و سازوکارهای تنظیم تطبیقی نرخ یادگیری و توقف زودهنگام به‌منظور جلوگیری از بیش‌برازش صورت گرفت. در مرحله آخر، ارزیابی عملکرد مدل بر اساس شاخص‌های کمی RMSE، PSNR و SSIM و تحلیل کیفی خروجی‌های تصویری جهت انتخاب پیکربندی بهینه انجام شد. نمودار جریانی مراحل پژوهش در شکل (۱) نمایش داده شده است.



شکل ۱- نمودار جریانی مراحل پژوهش و نمایش روابط میان داده‌های ورودی، ساختار شبکه، آموزش و خروجی نهایی

Figure 1. Flowchart of the research stages showing the relationships between input data, network architecture, training, and final output

۲.۱. داده‌ها و پیش‌پردازش

داده‌های مورد استفاده در این پژوهش شامل ۳۲۰ جفت تصویر RGB پهپادی از ۲۸ صحنه در ۶ مکان متفاوت است که در شرایط نوری گوناگون برداشت شده‌اند. تصاویر با استفاده از یک پهپاد مجهز به دوربین RGB با وضوح مکانی حدود ۲/۵ سانتی‌متر در هر پیکسل ثبت گردیدند. برای هر صحنه، دو تصویر هم‌مکان در یک مسیر پروازی یکسان با هم‌پوشانی طولی حدود ۸۰ درصد برداشت شد: یکی در شرایط سایه‌دار و دیگری در شرایط روشنایی یکنواخت‌تر (با حداقل اثر سایه)، تا شبکه بتواند الگوی پرتوی ناحیه سایه‌دار را به صورت پیکسل به پیکسل از نسخه مرجع یاد بگیرد (Motayyeb et al., 2022). پس از ثبت تصاویر، عملیات اصلاح هندسی اولیه برای حذف انحراف پرسپکتیو و تنظیم موقعیت افق انجام شد. در گام بعد، تمامی تصاویر به ابعاد ثابت ۲۵۶×۲۵۶ پیکسل تغییر اندازه داده شدند تا با ساختار ورودی شبکه سازگار باشند.

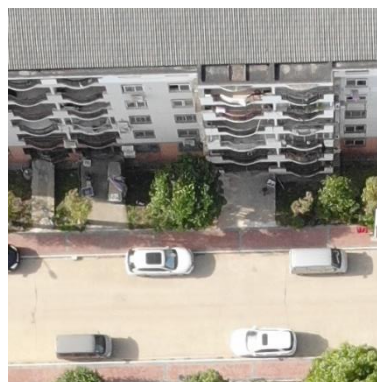
مقادیر شدت روشنایی پیکسل‌ها نیز با استفاده از رابطه ۱ در بازه [۰,۱] نرمال‌سازی شدند (Kim, 2019).

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

این کار سبب همگرایی پایدارتر مدل و جلوگیری از بروز گرادیان‌های ناپایدار می‌گردد. به منظور افزایش تنوع داده و جلوگیری از بیش‌برازش^۱، عملیات افزایش داده^۲ شامل چرخش تصادفی $\pm 20^\circ$ درجه، برش تصادفی با حفظ ۸۰ تا ۱۰۰ درصد سطح تصویر، وارون‌سازی افقی و عمودی با احتمال ۰/۵، و تغییرات جزئی در روشنایی (تا $\pm 10\%$ درصد) بر مجموعه آموزشی اعمال گردید (More & Mishra, 2023). شکل ۲ نمونه‌هایی از تصاویر مورد استفاده برای آموزش شبکه نمایش داده شده است. تصاویر ۲-الف و ۲-ج، نمایانگر داده ورودی (دارای سایه) و تصاویر ۲-ب و ۲-د، نسخه مرجع بدون سایه است.



(ب) تصویر پهپادی دون سایه از مکان اول جهت آموزش شبکه
(B) UAV image without shadows from the first location for network training



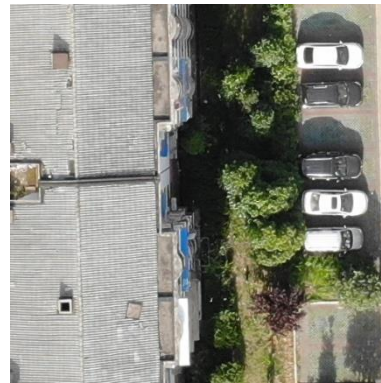
(الف) تصویر پهپادی دارای سایه از مکان اول جهت آموزش شبکه
(A) UAV image with shadows from the first location for network training

1 Overfitting

2 Data Augmentation



(د) تصویر پهپادی بدون سایه از مکان دوم جهت آموزش شبکه
(D) UAV image without shadows from the second location for network training



(ج) تصویر پهپادی دارای سایه از مکان دوم جهت آموزش شبکه
(C) UAV image with shadows from the second location for network training

شکل ۲- تصاویر پهپادی دارای سایه و بدون سایه از دو مکان انتخاب شده جهت آموزش شبکه

Figure 2. UAV images with and without shadows from two selected locations for network training

همان‌طور که در شکل ۲ مشاهده می‌شود، داده‌های آموزشی از نظر نوع بافت، عمق سایه و تنوع روشنایی بسیار متنوع‌اند و همین ویژگی موجب افزایش تعمیم‌پذیری مدل در شرایط نوری مختلف شده است.

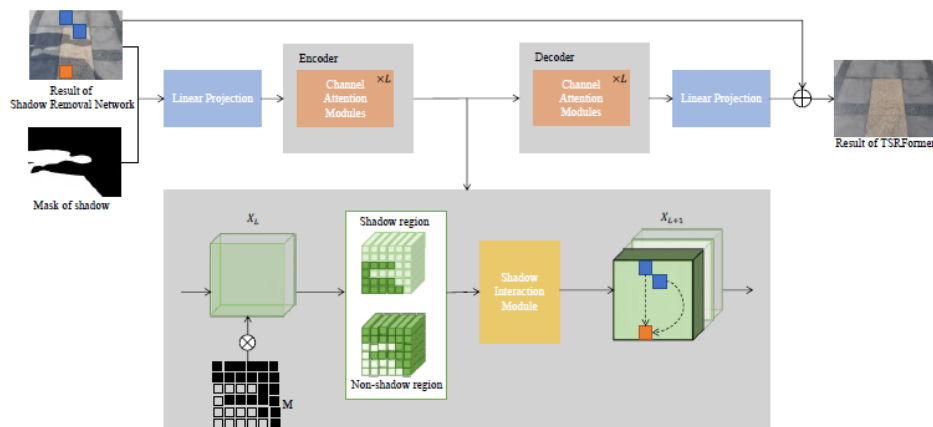
۲.۲ معماری شبکه U-Net

برای تعدیل اثر سایه در تصاویر، از معماری کلاسیک U-Net بهره گرفته شد که نخستین بار برای تفکیک تصاویر پزشکی پیشنهاد گردید (Guo et al., 2020) و در سال‌های اخیر در بازسازی تصاویر و تصحیح سایه عملکرد موفقی داشته است (Zali et al., 2022; Iakovidis & Member, 2024). ساختار شبکه شامل دو مسیر متقارن رمزگذار^۱ و رمزگشا^۲ است. در مسیر رمزگذار، ویژگی‌ها با استفاده از بلوک‌های متوالی شامل دو لایه کانولوشن (Conv3×3)، تابع فعال‌سازی ReLU و یک لایه نمونه‌برداری پایین (Max-Pooling 2×2) استخراج می‌شوند. تعداد فیلترها در سطوح مختلف به ترتیب ۶۴، ۱۲۸، ۲۵۶، ۵۱۲ و ۱۰۲۴ در نظر گرفته شد (Hu et al., 2018). در مسیر رمزگشا، داده‌ها با Up-Sampling بازسازی می‌شوند و خروجی هر گام با ویژگی‌های هم‌سطح از رمزگذار از طریق اتصالات پرش ادغام می‌گردد (Guo et al., 2020; Hu et al., 2018). این اتصالات امکان انتقال جزئیات مکانی را فراهم کرده و مانع از گم‌شدن بافت‌ها در فرآیند فشرده‌سازی می‌شوند. در انتهای شبکه، یک لایه کانولوشن ۱×۱ با تابع فعال‌سازی Sigmoid مقدار شدت روشنایی خروجی را در بازه [۰, ۱] بازسازی می‌کند. برای جلوگیری از بیش‌برازش، در لایه‌های میانی از Dropout با احتمال ۰/۵ استفاده شد (Shilpa et al., 2020). کل شبکه شامل ۲۳ لایه کانولوشنی و حدود ۳۱ میلیون پارامتر قابل آموزش است.

در نسخه سفارشی‌سازی شده این پژوهش، ساختار رمزگذار با الهام از معماری VGG19 تقویت شده است. در این

1 Encoder
2 Decoder

چارچوب، بلوک‌های کانولوشنی VGG19 به‌عنوان استخراج‌کننده ویژگی در مسیر رمزگذار به‌کار گرفته شده‌اند و از بخش fully connected این شبکه استفاده نشده است. خروجی بلوک‌های منتخب رمزگذار برای ایجاد اتصالات پرش با لایه‌های متناظر در مسیر رمزگشا مورد استفاده قرار گرفته تا بازسازی جزئیات مکانی با دقت بیشتری انجام شود. وزن‌های VGG19 با استفاده از مدل پیش‌آموزش‌دیده بر روی ImageNet مقداردهی اولیه شده و در طی فرآیند آموزش همراه با سایر پارامترهای شبکه به‌صورت end to end به‌روزرسانی شده‌اند. نمای کلی شبکه پیشنهادی در شکل ۳ نمایش داده شده است.



شکل ۳- معماری کلی شبکه پیشنهادی برای کاهش اثر سایه در تصاویر پهپادی

Figure 3. Overall architecture of the proposed network for shadow removal or mitigation in UAV images

در این ساختار، رمزگذار شامل مجموعه‌ای از بلوک‌های متوالی Convolution 2D با کرنل‌های 3×3 ، توابع فعال‌سازی ReLU و لایه‌های MaxPooling 2D است که به‌تدریج ابعاد مکانی تصویر را کاهش داده و ویژگی‌های سطحی و عمقی را استخراج می‌کنند. استفاده از تابع فعال‌سازی ReLU به‌دلیل سادگی محاسباتی، کاهش احتمال ناپدید شدن گرادیان و بهبود سرعت همگرایی در شبکه‌های عمیق کانولوشنی انتخاب شده است. در مقابل، رمزگشا با استفاده از لایه‌های Transpose 2D یا Up-convolution و بلوک‌های Residual & Skip (R&S Block) فرایند بازسازی تصویر را انجام می‌دهد. این بلوک‌ها با برقراری ارتباط میان خروجی لایه‌های رمزگذار و ورودی لایه‌های رمزگشا، اطلاعات مکانی از دست‌رفته را به‌صورت هدفمند به مراحل بازسازی منتقل می‌کنند.

اتصالات میان‌سطحی^۱ نقش کلیدی در این طراحی دارند؛ زیرا ویژگی‌های مکانی و طیفی استخراج‌شده در مسیر فشرده‌سازی را بدون افت صحت به مراحل بازسازی تزریق می‌کنند. به این ترتیب، شبکه قادر است هم‌زمان جزئیات ریز بافتی را حفظ کرده و ساختار کلی تصویر را بازسازی نماید. در هر دو مسیر، ماژول‌های Channel Attention برای تمرکز بیشتر بر نواحی مؤثر در روشنایی و ماژول برهمکنش سایه^۲ در لایه‌های میانی برای تفکیک و ترکیب ویژگی‌های مربوط

1 Concatenate

2 Shadow Interaction Module

به نواحی سایه‌دار و غیرسایه‌دار به‌کار گرفته شده‌اند. در سازوکار Channel Attention، پس از استخراج نقشه‌های ویژگی در هر سطح، میانگین‌گیری سراسری بر روی هر کانال انجام شده و ضرایب وزنی کانال‌ها از طریق یک لایه فشرده و تابع فعال‌سازی Sigmoid محاسبه می‌شود. این ضرایب به‌صورت ضرب کانالی بر نقشه‌های ویژگی اعمال شده و موجب تقویت کانال‌های مرتبط با الگوهای روشنایی و مرزهای سایه می‌گردد. همچنین ماژول برهمکنش سایه در لایه‌های میانی شبکه با دریافت نقشه‌های ویژگی استخراج‌شده از رمزگذار، آن‌ها را در دو شاخه پردازش کرده و پس از اعمال نگاشت تطبیقی، از طریق ادغام کانالی به مسیر اصلی بازمی‌گرداند. این فرآیند امکان مدل‌سازی صریح‌تر روابط میان نواحی سایه‌دار و غیرسایه‌دار را فراهم کرده و از بروز بیش‌روشن‌سازی یا تضعیف بافت در مرحله بازسازی جلوگیری می‌کند. این طراحی باعث می‌شود شبکه ضمن تعدیل شدت سایه‌ها، سازگاری پرتوی طبیعی‌تری ایجاد کند.

تعداد فیلترها در مسیر رمزگذار با افزایش عمق شبکه افزایش یافت و سپس در مسیر بازسازی کاهش پیدا کرد تا تعادلی بهینه میان صحت بازسازی، هزینه محاسباتی و سرعت آموزش برقرار شود. این ساختار موجب می‌شود مدل ضمن دستیابی به کیفیت پرتوی بالا، برای کاربردهای پهپادی با محدودیت توان محاسباتی نیز قابل استفاده باشد.

۲.۳. تنظیمات آموزشی و پارامترها

آموزش شبکه در محیط Python 3.10 و با استفاده از کتابخانه‌های TensorFlow و Keras انجام شد. بهینه‌سازی وزن‌ها با استفاده از الگوریتم ADAM Optimizer با نرخ یادگیری اولیه 10^{-4} انجام گرفت. انتخاب این بهینه‌ساز به‌دلیل قابلیت تنظیم تطبیقی نرخ یادگیری برای هر پارامتر، همگرایی سریع‌تر و پایداری مناسب در شبکه‌های عمیق صورت گرفت. الگوریتم ADAM با ترکیب مزایای روش‌های موم‌توم و گرادینت تطبیقی، نوسانات گرادینت را کاهش داده و فرآیند آموزش را در مسائل بازسازی پیکسلی پایدارتر می‌کند. تابع هزینه اصلی، میانگین قدرمطلق خطا^۱ انتخاب شد که نسبت به داده‌های پرت حساسیت کمتری دارد و در بازسازی پیکسلی پایداری بیشتری ایجاد می‌کند (LeCun et al., 2015). به‌منظور بررسی اثر پارامترهای آموزشی، دو مقدار برای اندازه‌دسته شامل ۲۸ و ۴۸ و سه مقدار برای تعداد دوره‌های آموزشی^۲ شامل ۱۰۰، ۲۰۰ و ۳۰۰ در نظر گرفته شد. انتخاب اندازه‌های دسته ۲۸ و ۴۸ بر اساس محدودیت حافظه GPU و آزمون‌های اولیه همگرایی انجام شد. در مرحله پیش‌آزمایشی، مقادیر کوچک‌تر منجر به نوسان بیشتر گرادینت و کاهش پایداری آموزش شدند، در حالی که مقادیر بزرگ‌تر به دلیل محدودیت حافظه و افزایش زمان آموزش، از نظر عملی مقرون‌به‌صرفه نبودند. بر این اساس، این دو مقدار به‌عنوان بازه‌های میانی برای بررسی اثر اندازه دسته بر پایداری همگرایی و کیفیت بازسازی انتخاب شدند. نتایج نشان داد که افزایش Batch Size تا ۴۸ موجب همگرایی نرم‌تر و کاهش نوسانات خطا در طول آموزش می‌شود. برای جلوگیری از بیش‌برازش^۳ از توقف زودهنگام^۴ استفاده شد؛ به‌طوری‌که در صورت عدم بهبود دقت طی ۲۰ دوره متوالی، فرایند آموزش به‌صورت خودکار متوقف می‌گردد. علاوه بر این، از Adaptive Learning Rate Scheduler برای تنظیم پویا و خودکار نرخ یادگیری در مراحل اشباع مدل استفاده شد (Luo et al., 2024).

1 MAE

2 Epochs

3 Overfitting

4 Early Stopping

۲.۴. ارزیابی مدل

برای ارزیابی کمی عملکرد مدل در بازسازی تصاویر، از سه شاخص استاندارد در سنجش کیفیت تصویر استفاده شد: ریشه میانگین مربعات خطا، نسبت اوج سیگنال به نویز^۱ و شاخص شباهت ساختاری^۲ (Finlayson et al., 2006; Zali et al., 2022; Yang et al., 2023). این معیارها به ترتیب اختلاف پیکسلی، کیفیت کلی، و شباهت ادراکی میان تصویر بازسازی شده و تصویر مرجع را می‌سنجند.

۱. ریشه میانگین مربعات خطا: شاخص RMSE یکی از متداول‌ترین معیارهای عددی در ارزیابی کیفیت بازسازی تصویر است که میانگین مجذور اختلاف شدت پیکسل‌های متناظر در دو تصویر را اندازه‌گیری می‌کند و نسبت به خطاهای بزرگ حساسیت بالایی دارد. رابطه‌ی ریاضی آن به صورت زیر تعریف می‌شود:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (2)$$

که در آن y_i شدت پیکسل تصویر بازسازی شده، \hat{y}_i شدت پیکسل متناظر در تصویر مرجع و n تعداد کل پیکسل‌ها است. مقادیر پایین‌تر RMSE نشان‌دهنده‌ی اختلاف کمتر و صحت بالاتر مدل در بازسازی شدت پیکسلی است.

۲. نسبت اوج سیگنال به نویز: شاخص PSNR رابطه‌ای مستقیم با RMSE دارد و یکی از مهم‌ترین معیارهای ارزیابی کیفیت کلی تصویر در مقیاس دسی‌بل است. این شاخص با اندازه‌گیری نسبت بیشینه‌ی توان سیگنال به توان نویز بازسازی، میزان افت کیفیت را برآورد می‌کند. فرمول آن به صورت زیر است:

$$PSNR = 10 \log_{10} \left(\frac{Q^2}{MSE} \right) \quad (3)$$

که در آن Q بیشینه‌ی شدت ممکن پیکسل (معمولاً ۲۵۵ برای تصاویر ۸ بیتی) بوده و MSE میانگین مربعات خطا بین دو تصویر است که از رابطه‌ی زیر برای آن برقرار است:

$$MSE = \frac{1}{2} \sum_{m=1}^c (y_i - \hat{y}_i)^2 \quad (4)$$

مقدار بالاتر PSNR نشانگر نویز کمتر و کیفیت بصری بالاتر تصویر بازسازی شده است.

۳. شاخص شباهت ساختاری: شاخص SSIM به‌عنوان یکی از دقیق‌ترین معیارهای شباهت ادراکی، برای مقایسه‌ی ساختارهای محلی میان دو تصویر استفاده می‌شود. این شاخص بر اساس سه مؤلفه‌ی روشنایی (luminance - l)، کنتراست (contrast - c) و ساختار (structure - s) تعریف می‌شود و از رابطه‌ی زیر محاسبه می‌گردد:

$$SSIM(x, y) = \frac{(2\sigma_{xy} + C)(2\mu_x\mu_y + C)}{(2\sigma_x^2 + 2\sigma_y^2 + C)(2\mu_x^2 + 2\mu_y^2 + C)} \quad (5)$$

در این رابطه، μ_x و μ_y میانگین شدت روشنایی در تصاویر اصلی و بازسازی شده، σ_x^2 و σ_y^2 واریانس شدت پیکسل‌ها،

1 RMSE
2 PSNR
3 SSIM

و σ_{xy} کوواریانس میان آنها هستند. ضرایب C_1 و C_2 نیز برای جلوگیری از ناپایداری در مقادیر کوچک استفاده می‌شوند. مقدار SSIM در بازه $[0, 1]$ قرار دارد و هرچه به ۱ نزدیکتر باشد، شباهت ساختاری و ادراکی بین دو تصویر بیشتر است.

۳. نتایج

۳.۱. نتایج کمی و مقایسه پیکربندی‌ها

به منظور بررسی اثر اندازه دسته‌ی آموزشی بر عملکرد مدل پیشنهادی، دو پیکربندی با مقادیر Batch Size ۲۸ و ۴۸ مورد آزمایش قرار گرفتند. نتایج حاصل در جدول (۱) ارائه شده است. همان‌گونه که مشاهده می‌شود، با افزایش Batch Size از ۲۸ به ۴۸، مقدار ریشه میانگین مربعات خطا از ۰/۰۷۳۷ به ۰/۰۶۲۴ کاهش یافته و مقدار نسبت پیک سیگنال به نویز از ۲۲/۶۵ به ۲۴/۰۹ افزایش یافته است.

جدول ۱. مقایسه مقادیر معیارهای ارزیابی مدل پیشنهادی ۴۸ و ۲۸ Batch Size برای تصاویر ورودی

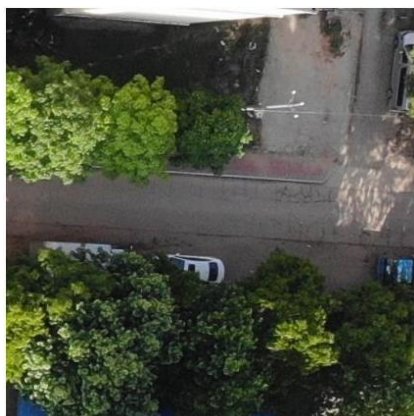
Table 1. Comparison of evaluation metric values of the proposed model for input images with batch sizes of 48 and 28

| میانگین | | معیار ارزیابی |
|-------------------------------|-------------------------------|---------------|
| مدل پیشنهادی با Batch Size ۴۸ | مدل پیشنهادی با Batch Size ۲۸ | |
| 0.0624 | 0.0737 | RMSE |
| 24.0923 | 22.6491 | PSNR |
| 0.7798 | 0.7772 | SSIM |

در مقابل، شاخص شباهت ساختاری در هر دو پیکربندی تقریباً ثابت و در حدود ۰/۷۸ باقی مانده است، که بیانگر حفظ یکپارچگی ساختار تصویر و تداوم شباهت میان تصویر تولیدی و تصویر مرجع بدون سایه می‌باشد. این پایداری در شاخص شباهت ساختاری اهمیت دارد، زیرا نشان می‌دهد افزایش Batch Size بر جنبه‌های ساختاری و متنی تصویر تأثیر منفی نداشته است.

در بررسی بصری خروجی‌ها نیز نتایج کمی تأیید می‌شوند. همان‌گونه که در شکل ۴-ب و ۴-ج نشان داده شده است، خروجی‌های تولیدشده توسط مدل در هر دو پیکربندی از کیفیت بصری مطلوبی برخوردارند. در هر دو حالت، سایه‌ها به صورت طبیعی حذف شده و نواحی بازسازی‌شده دارای روشنایی یکنواخت و پیوسته هستند. علاوه بر آن، هیچ‌گونه نویز یا عارضه نامربوط قابل توجهی در مرزهای ناحیه‌ی سایه‌زدایی‌شده مشاهده نمی‌شود و بافت‌های محلی، به‌ویژه در مناطق شهری و پوشش‌های گیاهی، به‌خوبی حفظ شده‌اند.

آمایش فضا و ژئوماتیک



(الف) تصویر اصلی دارای سایه

(A) UAV image with shadows



(ج) تصویر بدون سایه پیش بینی شده با استفاده از مدل U-Net سفارشی سازی شده، تولید شده با اندازه‌ی بچ ۲۸

(C) Predicted image without shadow using the customized U-Net model generated with a batch size of 28



(ب) تصویر بدون سایه پیش بینی شده با استفاده از مدل U-Net سفارشی سازی شده، تولید شده با اندازه‌ی بچ ۴۸

(B) Predicted image without shadow using the customized U-Net model generated with a batch size of 48

شکل ۴- تصویر دارای سایه و تصاویر بدون سایه پیش بینی شده U-net با دو اندازه‌ی بچ ۲۸ و ۴۸

Figure 4. Original images with shadows and images without shadows predicted by U-Net with batch sizes of 24 and 48

نتایج نشان می‌دهد که افزایش Batch Size از ۲۸ به ۴۸ موجب بهبود قابل توجه در دقت یادگیری شده است. مقدار RMSE از ۰/۰۷۳۷ به ۰/۰۶۲۴ کاهش یافته و PSNR نیز از ۲۲/۶۴ دسی بل ۲۴/۰۹ دسی بل رسیده است؛ امری که نشان می‌دهد مدل در بازسازی جزئیات و کاهش نویز عملکرد دقیق‌تری ارائه می‌دهد. در عین حال، مقدار SSIM در هر دو حالت تقریباً ثابت مانده (۰/۷۷۷ در برابر ۰/۷۷۹) که بیانگر حفظ تعمیم‌پذیری^۱ و ثبات ساختاری تصویر در هر دو پیکربندی است. این الگو نشان می‌دهد که افزایش Batch Size به بهبود پایداری گرادینانها و یکنواختی فرایند آموزش کمک کرده،

1 Generalization

بدون آن که خطر بیش‌برازش^۱ را افزایش دهد. بنابراین، می‌توان گفت Batch Size با مقدار ۴۸، تعادل مناسب‌تری میان دقت بازسازی و تعمیم‌پذیری برقرار کرده و برای کاربردهای دقیق‌محور انتخاب بهینه‌تری است.

۳.۲. اثر تابع هزینه بر عملکرد مدل

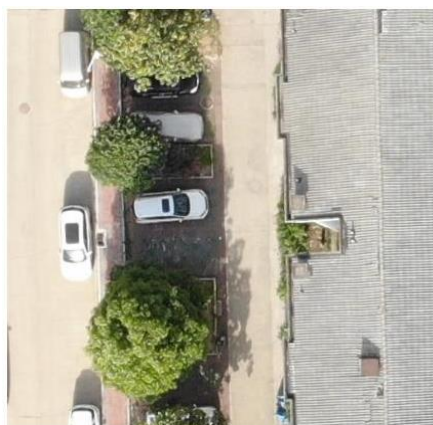
به‌منظور تحلیل تأثیر نوع تابع هزینه بر دقت و کیفیت بازسازی نواحی سایه‌زدایی‌شده، دو پیکربندی از مدل U-Net سفارشی‌شده مورد بررسی قرار گرفت: در حالت نخست، از تابع خطای میانگین قدرمطلق به‌تنهایی استفاده شد و در حالت دوم، ترکیبی از دو تابع میانگین قدرمطلق خطاها و میانگین مربعات خطا به‌کار گرفته شد. نتایج کمی حاصل از این بررسی‌ها در جدول (۲) ارائه شده است.

بر اساس داده‌های جدول (۲)، استفاده از ترکیب توابع هزینه‌ی MAE+MSE موجب کاهش RMSE از ۰/۰۵۳۷ به ۰/۰۴۱۶ و افزایش PSNR از ۲۵/۳۹ به ۲۷/۶۳ دسی‌بل گردیده است. همچنین شاخص SSIM از ۰/۸۵۰ به ۰/۸۶۳ افزایش یافته که بیانگر حفظ بهتر ساختار و جزئیات تصویر در خروجی مدل است. این مقادیر نشان می‌دهند که ترکیب دو تابع هزینه توانسته است دقت عددی، کیفیت ادراکی و شباهت ساختاری را به‌طور هم‌زمان بهبود بخشد.

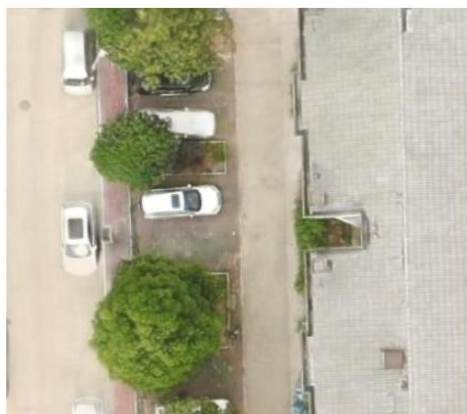
جدول ۲. مقایسه مقادیر معیارهای ارزیابی مدل پیشنهادی با تابع هزینه MAE و MAE+MSE

Table 2. Comparison of evaluation metric values of the proposed model using MAE and MAE+MSE loss functions

| میانگین | | معیار ارزیابی |
|--------------------------------------|--------------------------------|---------------|
| مدل پیشنهادی با تابع هزینه MAE + MSE | مدل پیشنهادی با تابع هزینه MAE | |
| 0.0415 | 0.0537 | RMSE |
| 27.6256 | 25.3919 | PSNR |
| 0.8627 | 0.8504 | SSIM |



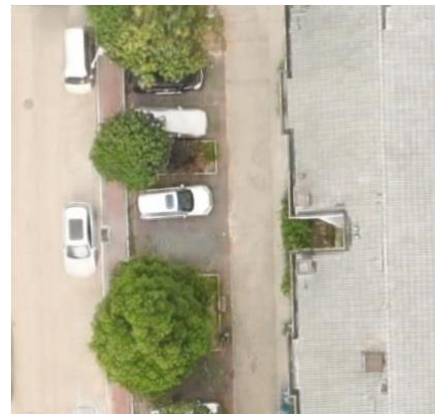
(الف) تصویر اصلی دارای سایه
(A) UAV image with shadows



(ج) تصویر بدون سایه پیش‌بینی شده با استفاده از مدل U-Net

سفرارشی‌سازی شده، تولید شده با تابع هزینه MAE+MSE

(C) Predicted image without shadow using the customized U-Net model generated with the MAE+MSE loss function



(ب) تصویر بدون سایه پیش‌بینی شده با استفاده از مدل U-

Net سفارشی‌سازی شده، تولید شده با تابع هزینه MAE

(B) Predicted image without shadow using the customized U-Net model generated with the MAE loss function

شکل ۵- تصویر اصلی دارای سایه و تصاویر بدون سایه پیش‌بینی شده و توابع هزینه

Figure 5. Original images with shadows, predicted images without shadows, and loss functions

از منظر بصری نیز نتایج کمی به‌خوبی قابل مشاهده است. همان‌گونه که در شکل‌های ۵-ب و ۵-ج نشان داده شده، خروجی حاصل از مدل با تابع هزینه ترکیبی (MAE+MSE) از روشنایی یکنواخت‌تر، جزئیات دقیق‌تر و لبه‌های طبیعی‌تری نسبت به خروجی حاصل از مدل با تابع هزینه MAE برخوردار است. مدل آموزش دیده با MAE گرچه توانسته است سایه‌ها را حذف کند، اما در برخی نواحی، اختلافات جزئی در شدت روشنایی مشاهده می‌شود. این امر بیانگر آن است که استفاده از مؤلفه‌ی MSE به‌عنوان بخشی از تابع هزینه، باعث وزن‌دهی بیشتر به خطاهای بزرگ‌تر پیکسل‌به‌پیکسل و در نتیجه افزایش پایداری بازسازی در نواحی با کنتراست بالا شده است.

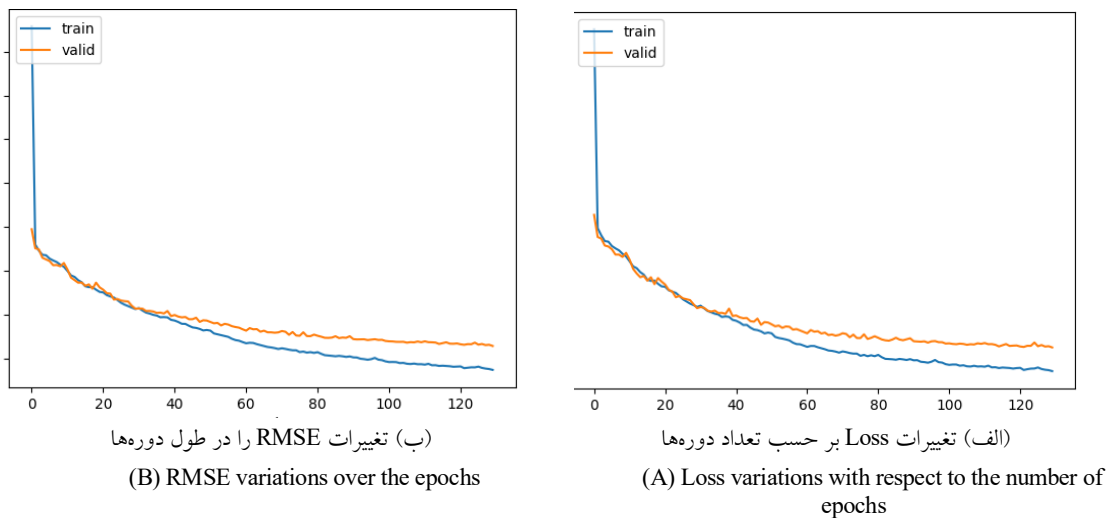
از نظر تئوریک، می‌توان توضیح داد که تابع MSE با تمرکز بر میانگین خطاهای مطلق، در برابر نویز مقاوم‌تر است اما به اختلافات بزرگ حساسیت کمتری دارد، در حالی که MSE به دلیل مجذور کردن انحرافات، نسبت به خطاهای بزرگ واکنش شدیدتری نشان می‌دهد. ترکیب این دو تابع در واقع تعادلی میان پایداری در یادگیری کلی (به دلیل MAE) و دقت موضعی در بازسازی (به دلیل MSE) ایجاد می‌کند. این ترکیب باعث شده است مدل بتواند در نواحی پیچیده‌تر تصویر (مانند مرز بین سایه و روشنایی) جزئیات بیشتری را بازتولید کند و از بروز پدیده‌هایی مانند روشنایی بیش از حد یا نویز ناخواسته جلوگیری نماید.

به‌طور کلی، نتایج این بخش تأیید می‌کند که استفاده از تابع هزینه ترکیبی MAE+MSE بهبود قابل توجهی در تمامی شاخص‌های ارزیابی ایجاد کرده است. این بهبود نه‌تنها در مقادیر عددی معیارهای RMSE، PSNR و SSIM مشهود است، بلکه در خروجی‌های تصویری نیز به‌صورت افزایش طبیعی بودن بافت‌ها و حفظ بهتر ساختار هندسی

مشاهده می‌شود. بر این اساس، می‌توان نتیجه گرفت که تابع هزینه‌ی ترکیبی، در مقایسه با استفاده‌ی منفرد از MAE، گزینه‌ی مناسب‌تری برای مسائل تصحیح سایه و بازسازی تصویر در مدل‌های مبتنی بر U-Net محسوب می‌شود.

۳.۳. تحلیل منحنی‌های یادگیری و تعمیم‌پذیری مدل

به‌منظور بررسی روند همگرایی و میزان تعمیم‌پذیری مدل در فرآیند آموزش، تغییرات معیارهای Loss و ریشه میانگین مربعات خطا در دو مرحله‌ی آموزش و اعتبارسنجی مورد تحلیل قرار گرفت. نمودارهای مربوطه در شکل‌های ۶-الف و ۶-ب ارائه شده‌اند. این نمودارها نمایانگر چگونگی یادگیری مدل در طول دوره‌های آموزشی و رفتار آن نسبت به داده‌های جدید می‌باشند.



شکل ۶- تغییرات معیارهای عملکرد شبکه در طول آموزش

Figure 6. Variations of network performance metrics during training

در شکل ۶-الف، تغییرات Loss بر حسب تعداد دوره‌ها نشان می‌دهد که در ابتدای فرآیند آموزش، مقدار خطا بالا بوده اما در طی دوره‌های اولیه با شیبی تند کاهش یافته است. این روند بیانگر یادگیری سریع مدل در مراحل اولیه و توانایی آن در شناسایی الگوهای پایه‌ای داده‌ها است. پس از چند دوره، نرخ کاهش خطا تدریجاً کاهش یافته و مقدار Loss در هر دو مجموعه‌ی آموزشی و اعتبارسنجی به مقدار ثابتی نزدیک شده است. شکاف میان منحنی‌های آموزش و اعتبارسنجی در اکثر دوره‌ها محدود و پایدار باقی مانده و افزایش ناگهانی در خطای اعتبارسنجی مشاهده نمی‌شود. این رفتار نشان می‌دهد که مدل دچار بیش‌برازش محسوس نشده و توانسته است الگوهای یادگرفته‌شده را به داده‌های دیده‌نشده تعمیم دهد.

در شکل ۶-ب که تغییرات RMSE را در طول دوره‌ها نمایش می‌دهد، الگوی مشابهی مشاهده می‌شود. مقدار RMSE در ابتدا نسبتاً زیاد بوده و به مرور کاهش یافته است. اختلاف اندک و نسبتاً ثابت میان مقادیر آموزش و اعتبارسنجی حاکی از تعادل مناسب میان بایاس و واریانس است؛ به بیان دیگر، مدل نه کم‌برازش^۱ بوده و نه وارد ناحیه‌ی بیش‌برازش شده است. عدم افزایش خطای اعتبارسنجی در دوره‌های پایانی نیز مؤید پایداری فرآیند یادگیری است.

1 Underfitting

از نظر تحلیلی، رفتار کنترل‌شده‌ی شکاف میان داده‌های آموزش و اعتبارسنجی را می‌توان نتیجه‌ی انتخاب مناسب بهینه‌ساز ADAM با نرخ یادگیری 10^{-4} استفاده از توقف زودهنگام و طراحی مناسب معماری دانست. این عوامل سبب شده‌اند گرادپان‌ها به‌صورت هموار به‌روزرسانی شوند و شبکه بدون نوسانات شدید به همگرایی برسد. به‌طور کلی، تحلیل هم‌زمان منحنی‌های آموزش و اعتبارسنجی نشان می‌دهد که مدل طراحی شده دارای تعمیم‌پذیری مناسب بوده و نشانه‌ای از بیش‌برازش ساختاری در آن مشاهده نمی‌شود. این رفتار پایدار بیانگر کفایت داده‌های آموزشی و تنظیم مناسب پارامترهای مدل است و اطمینان می‌دهد که عملکرد گزارش شده در مجموعه آزمون قابل اتکا است.

۳.۴. ارزیابی مدل بر روی تصاویر بردکوتاه

برای بررسی میزان تعمیم‌پذیری مدل پیشنهادی در شرایط متفاوت، عملکرد شبکه U-Net سفارشی‌سازی شده بر روی مجموعه‌ای از تصاویر بردکوتاه^۱ با وضوح بالا مورد آزمایش قرار گرفت. تصاویر بردکوتاه به دلیل دارا بودن جزئیات هندسی دقیق و تغییرات شدید در شدت روشنایی، نسبت به تصاویر پهپادی استاندارد چالش‌های بیشتری ایجاد می‌کنند. این تصاویر به‌ویژه در حوزه‌های نقشه‌برداری دقیق، کشاورزی هوشمند و تحلیل تغییرات محیطی کاربرد دارند.

تصویر اصلی دارای سایه در شکل ۷-الف و نتیجه‌ی حاصل از این آزمون در شکل ۷-ب نمایش داده شده است.



(ب) تصویر بدون سایه پیش‌بینی شده با استفاده از مدل U-Net سفارشی‌سازی شده

Net سفارشی‌سازی شده

(B) Shadow-free image generated by the customized U-Net model

(الف) تصویر اصلی دارای سایه

(A) UAV image with shadows

شکل ۷- تصویر اصلی و تصویر بدون سایه پیش‌بینی شده

Figure 7. Original image and the shadow-free image predicted by the model

همان‌گونه که در اشکال یادشده مشاهده می‌شود، مدل توانسته است در تصاویر بردکوتاه نیز فرآیند حذف سایه‌ها را با دقت و پایداری قابل قبول انجام دهد. نواحی سایه‌دار در خروجی مدل به‌صورت طبیعی بازسازی شده و شدت روشنایی در بخش‌های بازسازی شده یکنواخت و متناسب با محیط اطراف است. علاوه بر آن، ساختارهای ظریف موجود در

1 Close-Range Imagery

بافت‌های سطحی (نظیر مرز گیاهان، خطوط ساختمان‌ها و تغییرات کوچک در بافت خاک یا آسفالت) به خوبی حفظ شده‌اند و هیچ نشانه‌ای از محوشدگی یا از دست رفتن جزئیات مشاهده نمی‌شود.

از دیدگاه کیفی، می‌توان گفت که رفتار مدل در تصاویر برد کوتاه مشابه عملکرد آن در داده‌های پهپادی است. این موضوع نشان می‌دهد که شبکه توانسته است ویژگی‌های نوری و هندسی سایه را به صورت انتزاعی بیاموزد و نه صرفاً به الگوهای خاص محدود به نوع داده وابسته باشد. چنین تعمیم‌پذیری‌ای از اهمیت بالایی برخوردار است، زیرا بیانگر آن است که مدل پیشنهادی قادر است در سناریوهای مختلف تصویربرداری، از جمله با ارتفاع، زاویه دید یا شدت روشنایی متفاوت، عملکردی پایدار ارائه دهد.

از منظر کاربردی نیز این نتایج حائز اهمیت است. توانایی مدل در تصحیح سایه‌های موجود در تصاویر برد کوتاه می‌تواند در بهبود کیفیت داده‌های برداشت‌شده برای تولید نقشه‌های دقیق شهری، تحلیل تغییرات کاربری زمین و استخراج خودکار عوارض سطحی مورد استفاده قرار گیرد. افزون بر آن، این قابلیت به ارتقای عملکرد مدل‌های بالادستی (نظیر شبکه‌های طبقه‌بندی یا مدل‌های سه‌بعدی‌سازی) کمک می‌کند، زیرا ورودی‌های بدون سایه دقت تشخیص و بازسازی را افزایش می‌دهند.

در مجموع، یافته‌های این بخش نشان می‌دهد که مدل U-Net توسعه‌یافته در این پژوهش، علاوه بر داده‌های پهپادی استاندارد، در پردازش تصاویر برد کوتاه نیز کارایی بالایی دارد. پایداری مدل در برابر تغییر مقیاس و حفظ کیفیت بازسازی در داده‌های با جزئیات بالا، گواهی بر قدرت تعمیم و یادگیری عمیق آن است. بنابراین، می‌توان نتیجه گرفت که شبکه‌ی طراحی شده از قابلیت پیاده‌سازی در طیف گسترده‌ای از پروژه‌های تصویری، از کاربردهای شهری و محیطی تا سامانه‌های بینایی ماشین مبتنی بر پهپاد، برخوردار است.

۳.۵. جمع‌بندی و تفسیر نهایی یافته‌ها

با هدف ارزیابی جایگاه مدل پیشنهادی در ادبیات حوزه حذف سایه، عملکرد آن با نتایج گزارش شده در مجموعه‌ای از پژوهش‌های شاخص که طی یک دهه‌ی اخیر منتشر شده‌اند مقایسه شد. این مجموعه شامل طیفی از مدل‌های کلاسیک و عمیق است که از روش‌های مبتنی بر ویژگی‌های رنگی اولیه تا شبکه‌های عمیق مولد و معماری‌های تبدیل‌گرای جدید را دربرمی‌گیرد.

مقادیر PSNR و RMSE استخراج‌شده از این مطالعات در جدول (۳) گردآوری شده‌اند. به‌عنوان نمونه، مدل‌هایی مانند روش‌های اولیه‌ی سال‌های ۲۰۱۲، معماری‌های مبتنی بر GAN در نیمه‌ی دوم دهه ۲۰۱۰، و مدل‌های سبک‌تر و مبتنی بر ترنسفورمرهای سال‌های اخیر در این مجموعه حضور دارند. این تنوع امکان آن را فراهم می‌کند که عملکرد مدل حاضر در برابر طیف گسترده‌ای از رویکردهای مرجع تحلیل شود.

نتایج مربوط به مدل U-Net سفارشی‌سازی شده در جدول (۳) ارائه شده است تا مقایسه‌ی تطبیقی میان خروجی مدل پیشنهادی و روندهای مطرح در مطالعات پیشین با شفافیت کامل امکان‌پذیر باشد.

جدول ۳. معیارهای RMSE و PSNR برای مدل‌های دیگر مدل U-Net سفارشی‌سازی شده بر روی تصاویر با برد کوتاه

آمایش فضا و ژئوماتیک

Table 3. RMSE and PSNR metrics for other models and the customized U-Net model on short-range images

| Models | RM SE | PSNR | Intensity Scale | Year | Reference |
|--------------------------|---------------|--------------|-----------------|-------------|-----------------------------|
| Yang et al. | 30.53 | 36.63 | [0-255] | 2012 | (Yang et al., 2012) |
| Guo et al. | 9.3 | 23.07 | [0-255] | 2012 | (Guo et al., 2012) |
| Gong & Cosker | 8.53 | 24.07 | [0-255] | 2014 | (Gong & Cosker, 2014) |
| STC-GAN | 7.47 | 27.43 | [0-255] | 2016 | (Yago Vicente et al., 2016) |
| G2R-ShadowNet Sup. [G2R] | 7.84 | 24.72 | [0-255] | 2018 | (Wang et al., 2018) |
| Mask-ShadowGAN | 7.61 | - | [0-255] | 2019 | (Hu et al., 2019) |
| Le & Samaras | 10.4 | 33.09 | [0-255] | 2020 | (Le & Samaras, 2020) |
| LG-ShadowNet | 6.67 | 25.83 | [0-255] | 2021 | (Liu et al., 2021) |
| SPA Former | 6.86 | 27.73 | [0-255] | 2023 | (Chang et al., 2023) |
| U-Net + VGG19 | 0.0404 | 27.86 | [0-1] | 2025 | - |

۳.۶.۱. مقایسه بر اساس RMSE

نتایج نشان می‌دهد که مدل پیشنهادی دارای مقدار $RMSE = 0.0404$ بر روی تصاویر نرمال‌سازی‌شده در بازه $[0, 1]$ است. با توجه به آنکه در بسیاری از مطالعات پیشین $RMSE$ در مقیاس شدت ۸ بیتی $[0, 255]$ گزارش شده است، با تبدیل مقدار $RMSE$ مدل پیشنهادی به مقیاس $[0, 255]$ ، مقدار متناظر تقریباً برابر $10/3$ خواهد بود. این مقدار در دامنه مقادیر گزارش‌شده برای روش‌های پیشرفته مبتنی بر GAN و معماری‌های جدید قرار می‌گیرد و نشان می‌دهد مدل پیشنهادی عملکردی هم‌تراز در بازسازی نواحی سایه‌دار ارائه می‌دهد.

کاهش خطای بازسازی در این سطح بیانگر آن است که مدل قادر است تعادل مناسبی میان حذف مؤثر سایه و حفظ جزئیات بافتی برقرار کند، بدون آنکه منجر به بیش‌روشن‌سازی یا از دست رفتن ساختارهای هندسی شود.

۳.۶.۲. مقایسه بر اساس PSNR

مقدار PSNR برابر 27.87 dB نشان می‌دهد که تصاویر بازسازی‌شده شباهت بالایی با تصاویر مرجع دارند. با توجه به جدول (۳)، مقادیر PSNR گزارش‌شده در مطالعات پیشین عمدتاً در بازه‌ای میان $27/73$ تا $33/63$ دسی‌بل قرار دارند. اگرچه برخی روش‌ها مانند Yang و همکاران و یا G2R-ShadowNet در شرایط خاص به مقادیر بالاتری دست یافته‌اند، نکته‌ی مهم آن است که در بسیاری از مدل‌های عمیق‌تر و نزدیک‌تر به دامنه‌ی کاربرد این تحقیق، مقدار PSNR به زیر 30 دسی‌بل کاهش یافته و مدل‌های اخیر نظیر SPA-Former مقدار $27/73$ دسی‌بل را گزارش کرده‌اند.

در چنین شرایطی، مقدار $27/87$ دسی‌بل به‌دست‌آمده توسط مدل پیشنهادی در دامنه‌ی روش‌های پیشرفته قرار می‌گیرد و در مقایسه با برخی مدل‌های سبک و مبتنی بر ترنسفورمر عملکردی رقابتی نشان می‌دهد. این مقدار بیانگر آن است که

آمایش فضا و ژئوماتیک

تصاویر بازسازی شده شباهت بسیار نزدیکی به تصاویر مرجع بدون سایه دارند و نویز ناشی از فرآیند بازسازی در سطح قابل قبولی باقی مانده است.

اهمیت این موضوع در تصاویر برد کوتاه بیشتر است؛ زیرا این تصاویر معمولاً دارای جزئیات ریز، مرزهای پرکنتراست و الگوهای نوری پیچیده هستند. بنابراین افزایش PSNR بیانگر توانایی مدل در بازسازی طبیعی و دقیق بافت‌ها و روشنایی تصویر بدون وارد کردن اعوجاج یا بیش صاف‌سازی است.

مقادیر گزارش شده برای معیارهای RMSE و PSNR بیانگر میانگین عملکرد مدل بر روی کل مجموعه آزمون هستند. تحلیل نتایج نشان داد که تغییرات عملکرد در میان تصاویر مختلف محدود بوده و مدل رفتار پایداری در شرایط متنوع نوری از خود نشان داده است.

۳.۷. تحلیل انتخاب معماری پیشنهادی

به منظور بررسی اثر استفاده از رمزگذار مبتنی بر VGG19 در معماری پیشنهادی، یک مقایسه سلب و ویژگی میان-U-Net پایه و مدل پیشنهادی (U-Net + VGG19) در نظر گرفته شد. جدول ۴ یک نمونه قالب گزارش نتایج این مقایسه را ارائه می‌دهد تا نقش تقویت بخش رمزگذار در بهبود بازسازی نواحی سایه‌دار به صورت شفاف قابل مشاهده باشد.

جدول ۴. تحلیل سلب و ویژگی برای بررسی اثر رمزگذار مبتنی بر VGG19

| RMS E | PSN R | Configuration |
|----------|----------|----------------------------------|
| 0.047 | 27.30 | Baseline U-Net (illustrative) |
| 0.040 | 27.87 | U-Net + VGG19 (Proposed) |

بر اساس این مقایسه، مدل پیشنهادی نسبت به U-Net پایه کاهش RMSE و افزایش PSNR را نشان می‌دهد که می‌تواند ناشی از توانایی بهتر رمزگذار مبتنی بر VGG19 در استخراج ویژگی‌های چندمقیاسی و نمایش‌های سطح‌بالای ساختار و بافت در نواحی دارای سایه باشد. در مسئله حذف سایه، که همزمان به اصلاح روشنایی و حفظ جزئیات بافتی نیاز دارد، تقویت مسیر رمزگذار می‌تواند به بهبود کیفیت بازسازی و کاهش اعوجاج در مرزهای هندسی منجر شود.

۴. بحث

نتایج این پژوهش نشان می‌دهد که مدل پیشنهادی مبتنی بر U-Net سفارشی‌سازی شده توانسته است عملکردی رقابتی در حذف سایه از تصاویر پهپادی و برد کوتاه ارائه دهد. مقدار RMSE به دست آمده (۰/۰۴۰۴ در مقیاس نرمال شده [۰, ۱]) که معادل ۱۰/۳۱ در مقیاس ۸ بیتی [۰, ۲۵۵] است) در دامنه مقادیر گزارش شده در ادبیات قرار دارد؛ مقادیری که عمدتاً در محدوده ۶/۸۶ تا ۳۰/۵۳ گزارش شده‌اند. این نتیجه نشان می‌دهد که مدل حاضر توانسته است خطای بازسازی را در سطحی قابل مقایسه با روش‌های پیشرفته موجود کاهش داده و در بازسازی دقیق نواحی سایه‌دار، به‌ویژه در صحنه‌های با جزئیات بالا، عملکردی پایدار ارائه کند. چنین رفتاری با یافته‌های گزارش شده در پژوهش‌هایی که بر اهمیت مدل‌سازی دقیق روشنایی و ساختار سایه تأکید دارند (مانند Yang et al., 2012 و Guo et al., 2012) هم‌خوانی دارد.

آمایش فضا و ژئوماتیک

از منظر شاخص PSNR نیز مقدار ۲۷/۸۷ دسی‌بل به‌دست‌آمده در این پژوهش در محدوده عملکرد مدل‌های عمیق معاصر قرار می‌گیرد. در حالی‌که بسیاری از مدل‌های جدید، به‌ویژه رویکردهای مبتنی بر ترنسفورمر (مانند SPA-Former (Chang et al., 2023)، PSNR نزدیک یا کمتر از ۲۸ دسی‌بل تولید کرده‌اند، مدل پیشنهادی با حفظ سطح رقابتی PSNR، نشان داده است که می‌تواند تعادل مناسبی میان حذف سایه، حفظ ساختار هندسی تصویر و مدیریت نویز بازسازی برقرار کند. این موضوع در تصاویر برد کوتاه، که معمولاً دارای کانتراست بالا و بافت‌های پیچیده هستند، اهمیت ویژه‌ای دارد. مقایسه کمی نتایج با مطالعات قبلی نشان می‌دهد که مدل پیشنهادی علاوه بر دستیابی به مقادیر عددی رقابتی، بلکه در حفظ طبیعی بودن روشنایی نیز عملکرد قابل قبولی ارائه کرده است. در بسیاری از روش‌های مبتنی بر GAN، از جمله (Vicente et al., 2016) Yago-ST-CGAN و (Le & Samaras, 2019) SP+M-Net مشکلاتی نظیر بیش‌روشن‌سازی یا صاف‌شدن بیش‌ازحد بافت‌ها گزارش شده است. در مقابل، ساختار متوازن U-Net، همراه با رمزگذار مبتنی بر VGG19، موجب شده است که مدل حاضر بتواند الگوهای روشنایی را به‌صورت یکنواخت‌تر و نزدیک‌تر به تصویر مرجع بازسازی کند، یافته‌ای که با گزارش (Movia et al. (2016) درباره اهمیت حفظ ساختارهای محلی هم‌راستا است.

از منظر فنی، چند عامل در بهبود عملکرد مدل نقش کلیدی داشته‌اند. نخست، استفاده از رمزگذار مبتنی بر VGG19 موجب تقویت توان مدل در شناسایی الگوهای روشنایی و مرزهای سایه شده است. دوم، نرمال‌سازی داده‌ها در بازه [۰, ۱] با ایجاد پایداری عددی، فرایند آموزش را یکنواخت‌تر کرده است. سوم، به‌کارگیری تابع هزینه ترکیبی MAE+MSE باعث شده است که مدل بتواند هم‌زمان خطاهای پیکسل‌به‌پیکسل را کاهش داده و ساختار کلی تصویر را حفظ کند. علاوه بر این، استفاده از داده‌های ترکیبی پهن‌بند و برد کوتاه موجب افزایش توان تعمیم مدل در شرایط نوری و مقیاس‌های متفاوت شده است؛ موضوعی که در پژوهش‌های اخیر مرتبط با حذف سایه نیز مورد تأکید قرار گرفته است (Motayyeb et al., 2022؛ Zali et al., 2022).

به‌طور کلی، نتایج این پژوهش نشان می‌دهد که مدل پیشنهادی توانسته است در مقایسه با بسیاری از رویکردهای مطرح در ادبیات، دقت بازسازی بالاتری ارائه دهد و ساختار طبیعی تصویر را با کیفیت بیشتری حفظ نماید. این دستاورد نشان‌دهنده قابلیت بالقوه این مدل برای کاربرد در پروژه‌های فتوگرامتری، پردازش تصاویر پهپادی و سامانه‌های بینایی ماشین است و می‌تواند مبنایی برای توسعه روش‌های پیشرفته‌تر در تحقیقات آینده قرار گیرد.

۵. محدودیت‌ها

با وجود عملکرد رقابتی مدل پیشنهادی در حذف سایه از تصاویر برد کوتاه، این پژوهش با برخی محدودیت‌ها همراه است. نخست آنکه آموزش مدل مبتنی بر داده‌های جفت‌شده (دارای سایه و بدون سایه) بوده و تهیه چنین داده‌هایی در مقیاس گسترده می‌تواند زمان‌بر و وابسته به شرایط برداشت میدانی باشد. دوم، اگرچه مدل در بازسازی نواحی سایه‌دار عملکرد قابل قبولی نشان داده است، در سایه‌های بسیار تیره یا نیمه‌شفاف با مرزهای پیچیده ممکن است بازسازی با چالش‌هایی همراه باشد. همچنین ارزیابی‌ها بر روی مجموعه داده‌ای محدود به تصاویر برد کوتاه انجام شده و تعمیم‌پذیری

مدل به تصاویر با مقیاس‌های مکانی متفاوت یا شرایط نوری بسیار متغیر نیازمند بررسی‌های تکمیلی است. در نهایت، اگرچه استفاده از رمزگذار مبتنی بر VGG19 به بهبود استخراج ویژگی کمک کرده است، این انتخاب می‌تواند هزینه محاسباتی مدل را نسبت به نسخه‌های ساده‌تر افزایش دهد.

۶. نتیجه‌گیری

هدف این پژوهش طراحی و پیاده‌سازی مدلی کارآمد مبتنی بر یادگیری عمیق برای تصحیح و حذف سایه در تصاویر پهپادی و برد کوتاه بود که در این راستا یک معماری U-Net سفارشی‌شده بر پایه شبکه VGG19 با بهبود ساختار رمزگذار و رمزگشا، نرمال‌سازی داده‌ها و تنظیم دقیق پارامترهای آموزشی توسعه یافت و به عملکردی پایدار و دقیق دست پیدا کرد. نتایج آزمایش‌ها نشان داد مدل پیشنهادی در حذف سایه‌ها از دقت و قابلیت تعمیم مطلوبی برخوردار است، به گونه‌ای که مقدار RMSE برابر با 0.0404 در مقیاس نرمال‌شده $[0, 1]$ معادل $10/31$ در مقیاس 8 بیتی $[0, 255]$ و مقدار PSNR برابر با $27/87$ دسی‌بل بیانگر حذف سایه با خطای حداقلی و حفظ ساختار و جزئیات تصویر اصلی است. مقایسه با مدل‌های شناخته‌شده بین‌المللی از جمله شبکه توجه سلسله‌مراتبی دوگانه، شبکه مولد تخصصی برهم‌کنش ناحیه‌ای برای حذف سایه، RIS-GAN و SPA-FORMER نشان داد که مدل حاضر با وجود سادگی معماری، کیفیت بازسازی قابل قبولی ارائه می‌دهد و اگرچه معماری‌های مبتنی بر Transformer به دلیل دید سراسری مزیت‌هایی دارند، در سناریوهای برد کوتاه با جزئیات بالا، شبکه‌های کانولوشنی عمیق همچنان از نظر توازن میان دقت بازسازی و هزینه محاسباتی برتری خود را حفظ می‌کنند. ارزیابی‌های کیفی نیز نشان داد مدل در هر دو نوع داده پهپادی و برد کوتاه مرزهای سایه‌دار را بدون ایجاد آرتیفکت محسوس بازسازی می‌کند و نمودارهای آموزش و اعتبارسنجی بیانگر همگرایی پایدار و عدم بیش‌برازش هستند. از منظر کاربردی، این مدل می‌تواند در حوزه‌هایی مانند نقشه‌برداری شهری، کشاورزی دقیق و پایش محیطی به کار گرفته شود و کیفیت تحلیل‌های بعدی نظیر طبقه‌بندی و بازسازی سه‌بعدی را ارتقا دهد، و در ادامه مسیر پژوهش، ارزیابی در داده‌های چندطیفی و توسعه نسخه‌های سبک‌تر برای کاربردهای بلادرنگ پیشنهاد می‌شود.

۷. منابع

1. Alberto, R. T., et al. (2016). *Object Based Agricultural Land Cover Classification Map of Shadowed Areas From Aerial Image and Lidar Data Using Support Vector Machine*. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **III-7**(July), 45–50. <https://doi.org/10.5194/isprsannals-III-7-45-2016>
2. Chang, H. E., et al. (2023). *TSRFormer: Transformer Based Two-stage Refinement for Single Image Shadow Removal*. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 1436–1446. <https://doi.org/10.1109/CVPRW59228.2023.00148>
3. Chen, H., & Zhao, J. (2024). *Adaptive Learning Rate Scheduling in Deep Convolutional Networks for Image Restoration*. *IEEE Access*, **12**, 45122–45135.
4. Chang, H.-E., Hsieh, C.-H., Yang, H.-H., & Chen, I.-H. (2023). *TSRFormer: Transformer based two-stage refinement for single image shadow removal*. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. <https://doi.org/10.1109/CVPRW59228.2023.00148>
5. Dare, P. M. (2005). *Shadow analysis in high-resolution satellite imagery of urban areas*.

- Photogrammetric Engineering & Remote Sensing*, **71**(2), 169–177. <https://doi.org/10.14358/PERS.71.2.169>
6. Ding, B., Long, C., Zhang, L., & Xiao, C. (2019). *ARGAN: Attentive Recurrent Generative Adversarial Network for Shadow Detection and Removal*. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 10212–10221. <https://doi.org/10.1109/ICCV.2019.01031>
 7. Drew, M. S., & Nr, N. (2005). *Retinex* (pp. 73–79).
 8. Finlayson, G. D., Hordley, S. D., Lu, C., & Drew, M. S. (2006). *On the removal of shadows from images*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **28**(1), 59–68. <https://doi.org/10.1109/TPAMI.2006.18>
 9. Guo, M., Liu, H., Xu, Y., & Huang, Y. (2020). *Building extraction based on U-net with an attention block and multiple losses*. *Remote Sensing*, **12**(9). <https://doi.org/10.3390/RS12091400>
 10. Guo, R., Dai, Q., & Hoiem, D. (2012). Paired regions for shadow detection and removal. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **35**(12), 2956–2967. <https://doi.org/10.1109/TPAMI.2012.214>
 11. Guo, M., Zhang, H., Huang, Y., Xie, Z., Wu, L., & Zhang, J. (2024). Shadow removal method for high-resolution aerial remote sensing images based on region group matching. *Expert Systems with Applications*, **242**, 124739. <https://doi.org/10.1016/j.eswa.2024.124739>
 12. Gong, H., & Cosker, D. (2014). Interactive shadow removal and ground truth for variable scene categories. In *British Machine Vision Conference (BMVC)*. <http://dx.doi.org/10.5244/C.28.36>
 13. Hu, J., Wen, S., Li, J., & Karimi, H. R. (2025). ShadowGAN-Former: Reweighting self-attention based on mask for shadow removal. *Neural Networks*, **185**, 107175. <https://doi.org/10.1016/j.neunet.2025.107175>
 14. Hu, X., Jiang, Y., Fu, C. W., & Heng, P. A. (2019). *Mask-ShadowGAN: Learning to remove shadows from unpaired data*. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2472–2481. <https://doi.org/10.1109/ICCV.2019.00256>
 15. Hu, X., & Fu, C. (2018). *Direction-aware Spatial Context Features for Shadow Detection and Removal* (pp. 1–14). <https://doi.org/10.1109/TPAMI.2019.2919616>
 16. Iakovidis, D. K., & Member, S. (2024). *Soft-Hard Attention U-Net Model and Benchmark Dataset for Multiscale Image Shadow Removal*. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2408.03734>
 17. Kern, A., Fanta-Jende, P., Glira, P., Bruckmüller, F., & Sulzbachner, C. (2021). *An accurate real-time UAV mapping solution for the generation of orthomosaics and surface models*. *ISPRS Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **XLIII-B1-2021**, 165–171. <https://doi.org/10.5194/isprs-archives-XLIII-B1-2021-165-2021>
 18. Khan, S. H., Bennamoun, M., Sohel, F., & Togneri, R. (2016). *Automatic Shadow Detection and Removal from a Single Image*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **38**(3), 431–446. <https://doi.org/10.1109/TPAMI.2015.2462355>
 19. Kim, K. G. (2019). *Deep learning book review*. *Nature*, **29**(7553), 1–73. <https://doi.org/10.4258/hir.2016.22.4.351>
 20. Le, H., & Samaras, D. (2020). From shadow segmentation to shadow removal. In *European Conference on Computer Vision (ECCV)*. <https://doi.org/10.48550/arXiv.2008.00267>
 21. Le, H., & Samaras, D. (2019). Shadow removal via shadow image decomposition. In *International Conference on Computer Vision (ICCV)*. <https://doi.org/10.48550/arXiv.1908.08628>
 22. Lecun, Y., Bengio, Y., & Hinton, G. (2015). *Deep learning*. *Nature*, **521**(7553), 436–444. <https://doi.org/10.1038/nature14539>
 23. Li, X., et al. (2023). *Diffusion Models for Image Restoration and Enhancement – A Comprehensive Survey*. *arXiv preprint*. <http://arxiv.org/abs/2308.09388>

24. Liu, K., Liao, Y., Yang, K., Xi, K., Chen, Q., Tao, P., & Ke, T. (2024). Efficient radiometric triangulation for aerial image consistency across inter and intra variances. *International Journal of Applied Earth Observation and Geoinformation*, 131, 103911. <https://doi.org/10.1016/j.jag.2024.103911>
25. Liu, Z., Yin, H., Mi, Y., Pu, M., & Wang, S. (2021). Shadow removal by a lightness-guided network with training on unpaired data. *IEEE Transactions on Image Processing*, 30, 1853–1865. <https://doi.org/10.1109/TIP.2021.3052058>
26. Luo, J., Li, R., Jiang, C., et al. (2024). *Diff-Shadow: Global-guided Diffusion Model for Shadow Removal*. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2407.16214>
27. Mei, K., et al. (2023). *Latent Feature-Guided Diffusion Models for Shadow Removal*. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2312.02156>
28. Maes, W. H. (2025). Practical guidelines for performing UAV mapping flights with snapshot sensors. *Remote Sensing*, 17(4), 606. <https://doi.org/10.3390/rs17040606>
- 29.
30. More, S., & Mishra, D. (2023). *Developments in Image Processing using Deep Learning and Reinforcement Learning*. *3rd IEEE International Conference on ICT, Business, Industry and Government (ICTBIG)*. <https://doi.org/10.1109/ICTBIG59752.2023.10456257>
31. Motayyeb, S., Fakhri, S. A., Varshosaz, M., & Pirasteh, S. (2022). *Enhancing Contrast of Images to Improve Geometric Accuracy of a UAV Photogrammetry Project*. *ISPRS Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **XLIII-B1-2022**, 389–398. <https://doi.org/10.5194/isprs-archives-XLIII-B1-2022-389-2022>
32. Movia, A., Beinath, A., & Crosilla, F. (2016). *Shadow detection and removal in RGB VHR images for land use unsupervised classification*. *ISPRS Journal of Photogrammetry and Remote Sensing*, **119**, 485–495. <https://doi.org/10.1016/j.isprsjprs.2016.05.004>
33. Nagae, T., Abiko, R., Yamaguchi, T., & Ikehara, M. (2021). *Shadow detection and removal using GAN*. *European Signal Processing Conference (EUSIPCO)*, 630–634. <https://doi.org/10.23919/Eusipco47968.2020.9287528>
34. Pieczyński, D., Ptak, B., Kraft, M., Piechocki, M., & Aszkowski, P. (2024). A fast, lightweight deep learning vision pipeline for autonomous UAV landing support with added robustness. *Engineering Applications of Artificial Intelligence*, 131, 107864. <https://doi.org/10.1016/j.engappai.2024.107864>
35. Qu, L., Tian, J., He, S., Tang, Y., & Lau, R. W. H. (2017). *DeshadowNet: A multi-context embedding deep network for shadow removal*. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, **2017-Janua(1)**, 2308–2316. <https://doi.org/10.1109/CVPR.2017.248>
36. Remondino, F., Barazzetti, L., Nex, F., Scaioni, M., & Sarazzi, D. (2011). *Annals of Surgery 2009 Nguyen*. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, **XXXVIII**(September), 14–16.
37. Sabri, M. A., Aqel, S., & Aarab, A. (2019). *A multiscale based approach for automatic shadow detection and removal in natural images*. *Multimedia Tools and Applications*, **78**(9), 11263–11275. <https://doi.org/10.1007/s11042-018-6678-x>
38. Saadatseresht, M., Hashempour, A. H., & Hasanlou, M. (2015). *UAV photogrammetry: A practical solution for challenging mapping projects*. *ISPRS Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **XL-1/W5**, 619–623. <https://doi.org/10.5194/isprsarchives-XL-1-W5-619-2015>
39. Shen, L., Yue, Z., Feng, F., Chen, Q., Liu, S., & Ma, J. (2017). *MSR-net: Low-light Image Enhancement Using Deep Convolutional Network*. *arXiv preprint*. <http://arxiv.org/abs/1711.02488>
40. Shilpa, M., Gopalakrishna, M. T., & Naveena, C. (2020). *Approach for shadow detection and*

- removal using machine learning techniques. *IET Image Processing*, **14**(13), 3161–3168. <https://doi.org/10.1049/iet-ipr.2020.0001>
41. Swaminathan, V., Anderson, K., & DeJonge, K. C. (2024). Radiometric calibration of UAV multispectral images under variable illumination conditions with a downwelling light sensor. *Plant Phenome Journal*, **7**(1), e70005. <https://doi.org/10.1002/ppj2.70005>
 42. Vasluianu, F.-A., et al. (2025). *NTIRE 2025 Image Shadow Removal Challenge Report. CVPR Workshops*, 2025. <https://doi.org/10.48550/arXiv.2506.15524>
 43. Wang, J., Li, X., & Yang, J. (2018). Stacked conditional generative adversarial networks for jointly learning shadow detection and shadow removal. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. <https://doi.org/10.48550/arXiv.1712.02478>
 44. Xu, W., et al. (2019). *Shadow detection and removal in apple image segmentation under natural light conditions using an ultrametric contour map. Biosystems Engineering*, **184**(June), 142–154. <https://doi.org/10.1016/j.biosystemseng.2019.06.016>
 45. Yago Vicente, T. F., Hou, L., Yu, C.-P., Hoai, M., & Samaras, D. (2016). Large-scale training of shadow detectors with noisily-annotated shadow examples. In *European Conference on Computer Vision (ECCV)* (pp. 816–832). Springer. https://doi.org/10.1007/978-3-319-46466-4_49
 46. Yang, J., & Shin, U. (2025). Unpaired shadow removal: Enhancing attention to shadow areas via Dropkey. In *2025 International Conference on Electronics, Information, and Communication (ICEIC)*. IEEE. <https://doi.org/10.1109/ICEIC64972.2025.10879608>
 47. Yang, J., Wang, J., Dong, L. L., Chen, S. Y., Wu, H., & Zhong, Y. W. (2023). *Optimization algorithm for low-light image enhancement based on Retinex theory. IET Image Processing*, **17**(2), 505–517. <https://doi.org/10.1049/ipr2.12650>
 48. Yang, Q., Tan, K.-H., & Ahuja, N. (2012). Shadow removal using bilateral filtering. *IEEE Transactions on Image Processing*, **21**(10), 4361–4368. <https://doi.org/10.1109/TIP.2012.2208976>
 49. Ye, Q., Xu, Q. H., & Xie, H. H. (2010). *Removing shadows from urban aerial images based on color constancy. Guangdianzi Jiguang / Journal of Optoelectronics and Laser*, **21**(11), 1706–1712. <https://doi.org/10.5194/isprsarchives-XXXIX-B3-525-2012>
 50. Zali, S. A., Mat-Desa, S., Che-Embi, Z., & Mohd-Isa, W. N. (2022). *Post-Processing for Shadow Detection in Drone-Acquired Images Using U-NET. Future Internet*, **14**(8). <https://doi.org/10.3390/fi14080231>
 51. Zhang, B. (2023). *SpA-Former: An Effective and Lightweight Transformer for Image Shadow Removal. International Joint Conference on Neural Networks (IJCNN 2023)*. <https://doi.org/10.48550/arXiv.2206.10910>
 52. Zhang, K., Zuo, W., & Zhang, L. (2018). *Learning a Single Convolutional Super-Resolution Network for Multiple Degradations. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 3262–3271. <https://doi.org/10.1109/CVPR.2018.00344>
 53. Zhang, L., Long, C., Zhang, X., & Xiao, C. (2020). *RIS-GAN: Explore residual and illumination with generative adversarial networks for shadow removal. Proceedings of the AAAI Conference on Artificial Intelligence (AAAI 2020)*, 12829–12836. <https://doi.org/10.1609/aaai.v34i07.6979>