

شناسایی تغییرات تصاویر سنجش‌ازدوری با استفاده از روش‌های یادگیری عمیق دوجریان

یگانه حسن‌زاده^۱، عباس کیانی^{۲*}، نیما فرهادی^۳

^۱ دانشجوی کارشناسی - گروه مهندسی نقشه‌برداری - دانشگاه صنعتی نوشیروانی بابل

Yeganehasanzadeh3971@gmail.com

^۲ استادیار گروه مهندسی نقشه‌برداری - دانشکده عمران - دانشگاه صنعتی نوشیروانی بابل

a.kiani@nit.ac.ir

^۳ کارشناسی ارشد فتوگرامتری - دانشکده مهندسی نقشه‌برداری - دانشگاه صنعتی خواجه‌نصیرالدین طوسی

Farhadinima75@email.kntu.ac.ir

(تاریخ دریافت: خرداد ۱۴۰۲، تاریخ تصویب: شهریور ۱۴۰۲)

چکیده

شناسایی تغییرات تصاویر سنجش‌ازدور دارای نقش مهمی در بسیاری از برنامه‌های کاربردی مانند بررسی رشد شهرنشینی، نظارت بر تغییرات کاربری زمین، ارزیابی بلایا و آسیب‌های طبیعی ایفا می‌کند. در این فرایند هدف، تعیین برچسب تغییر کرده یا نکرده برای پیکسل‌های دو تصویر می‌باشد که از یک مکان ولی در دو زمان متفاوت گرفته شده‌است. همین‌طور در دهه اخیر استفاده از روش‌های یادگیری عمیق به دلیل عملکرد مناسبی که در تفسیر و پردازش داده‌های سنجش‌ازدوری دارند و همین‌طور توانایی حذف مهندسی ویژگی و استخراج ویژگی‌های سطح بالا از تصاویر، مورد توجه بسیاری از محققان این حوزه قرار گرفته است. در همین راستا، در این مقاله یک مدل یادگیری عمیق بهینه طراحی شده‌است که به دلیل ساختار سلسله‌مراتبی، طراحی مؤثر انتقال ویژگی و بهره‌وری مناسب از ویژگی‌های چندمقیاسه، دقت شناسایی تغییرات تصاویر دو زمانه را افزایش می‌دهد. مدل پیشنهادی به دلیل ساختار و معماری بهینه نسبت به برخی مدل‌های مشهور موجود مانند BIT از سرعت و دقت نتایج بالاتری برخوردار است. نتایج به دست آمده از اعمال مدل پیشنهادی بر روی دو مجموعه داده‌ی مورد بررسی، نسبت به روش مقایسه‌ای BIT به‌طور میانگین حاکی از دقت کلی حدود ۹۶ درصد و کاهش ۱۰ درصدی میزان محاسبات ضرب و جمع می‌باشد.

واژگان کلیدی: شناسایی تغییرات، سنجش از دور، یادگیری عمیق، تصاویر دوزمانه، شبکه‌های عصبی دوجریان

* نویسنده رابط

۱- مقدمه

هدف از شناسایی تغییرات تصاویر سنجش از دوری، تشخیص اختلافات به وجود آمده در یک محدوده‌ی جغرافیایی مشخص ولی در زمان‌های مختلف می‌باشد [۱] و محصول آن یک نقشه خواهد بود که بسته به نوع عارضه‌های مورد مطالعه متفاوت است. نقشه‌های تغییرات به دست آمده برای برنامه‌ریزان، مدیران و کارشناسان فنی حوزه‌های مختلف بسیار حائز اهمیت بوده و می‌توانند در مقیاس‌های کلان نیز مورد مطالعه قرار گیرند. با این حال، در مقیاس‌های بزرگ به علت دخیل شدن اقلیم‌های مختلف، شرایط محیطی در طیف وسیع‌تری دارای تغییرات است و میزان محاسبات و جزئیات مورد نیاز جهت تهیه نقشه تغییرات نیز افزایش می‌یابد [۲]. با در نظر گرفتن چنین مواردی، نیاز به روش‌هایی است که در عین محاسبات بهینه، توانایی مدیریت اختلافات موجود در تصاویر سنجش از دور را داشته باشند و نقشه‌ی تغییرات را با کیفیت بالاتری ارائه دهند.

روش‌هایی که جهت تهیه‌ی چنین نقشه‌هایی بکار می‌روند به طور معمول به دودسته‌ی سنتی و یادگیرنده تقسیم می‌شوند. روش‌های سنتی موجود در زمینه‌ی شناسایی تغییرات نیازمند افراد خبره در طراحی و انتخاب نوع ویژگی مورد استفاده است، از این جهت در مناطق مختلف عمومی‌سازی پایینی دارند [۳]. از طرفی، روش‌های یادگیرنده با حذف مهندسی ویژگی و استخراج ویژگی‌های سطح بالا [۴] در مقایسه با روش سنتی توانایی بالاتری در مدیریت اختلافات موجود در تصاویر دارند. طیف گسترده‌ای از تحقیقاتی که در زمینه‌ی شناسایی تغییرات یادگیرنده نظارتی با استفاده از شبکه‌های پیچشی انجام شده است را به دو دسته روش‌های دومرحله‌ای و روش‌های مجتمع (تک‌مرحله‌ای) می‌توان تقسیم کرد [۲، ۵]. در دسته اول، شبکه‌های پیچشی طبقه‌بندی تصویر زمان اول و دوم را انجام می‌دهند و پس از آن تصاویری که طبقه‌بندی شده‌اند برای تولید نقشه تغییرات نهایی، با یکدیگر مقایسه می‌شوند [۶-۸]. محدودیت این روش، نیازمند بودن به نقشه تغییرات و نقشه کلاسه‌بندی شده‌ی تصاویر دوزمانه است [۲]. در دسته دوم، نقشه تغییرات به صورت مستقیم از تصاویر دوزمانه محاسبه می‌شود. در بعضی تحقیقات [۹-۱۱] تصاویر دوزمانه به قطعه‌های

تصویری کوچک‌تری تقسیم می‌شوند و پس از اعمال مدل پیچشی، کلاس (تغییر کرده یا نکرده) مرکز قطعه تعیین می‌گردد. این دسته از روش‌ها به دلیل تعداد بالای دفعات اعمال مدل به تصویر، معمولاً میزان محاسبات بالایی دارند.

در راستای مقابله با محدودیت‌ها و هزینه‌ی محاسباتی بالای مدل‌های برپایه قطعه‌ی تصویری، تحقیقات بسیاری صورت گرفته‌است که عمده‌ی این تحقیقات به توسعه‌ی مدل‌های برپایه پیکسل تمرکز دارند. روش‌های برپایه پیکسل [۱، ۶، ۱۲-۲۵]، یک شبکه تماماً پیچشی به جفت تصویر اعمال می‌کنند و در نهایت نقشه تغییرات با یکبار اعمال مدل به دست می‌آید. برخی از روش‌های پیکسل پایه، تصویر زمان اول و دوم را در راستای باندها با یکدیگر ترکیب می‌کنند و بعد مدل پیچشی اعمال می‌گردد که به آن روش‌های پیش‌ترکیبی^۱ گفته می‌شود [۱]. از جمله تحقیقات در این زمینه می‌توان به [۲۶] اشاره کرد که تصاویر دوزمانه سنجش از دوری را پس از ترکیب از طریق باندها، وارد شبکه UNet++ [۲۷] بهبود یافته کردند و شناسایی تغییرات را در این تصاویر انجام دادند. دسته‌ای دیگر، مدل‌های پس-ترکیبی^۲ نامیده می‌شوند که در آن‌ها، استخراج ویژگی بر روی دو تصویر زمان اول و دوم انجام می‌گردد. سپس ویژگی‌های تولید شده با یکدیگر ترکیب و مقایسه می‌شوند و در نهایت نقشه تغییرات بر اساس ویژگی‌های ترکیب شده به دست می‌آید. به عنوان مثال، [۱۹] دو شبکه پس‌ترکیبی شناسایی تغییرات سنجش از دور را بر اساس معماری‌های دوگانه طراحی و با شبکه‌هایی از نوع روش‌های پیش‌ترکیبی مقایسه کردند. در این تحقیق ویژگی‌های دوزمانه‌ی استخراج شده در دو حالت ترکیبی و تفاضلی در کدگشا^۳ مورد استفاده قرار گرفتند. امکان ترکیب و تولید ویژگی‌های سطح بالاتر در چنین معماری‌هایی باعث افزایش بهره‌وری و عمومیت آن‌ها گردیده و توجه بسیاری از تحقیقات صورت گرفته در این زمینه را نیز به خود جلب کرده‌است.

بسیاری از محققان در راستای بهبود روش‌های پس ترکیبی، به دنبال استخراج و ترکیب ویژگی‌های چندمقیاسه بوده‌اند که از جمله مهم‌ترین آن‌ها می‌توان به Zhang و همکاران [۱] اشاره کرد. آن‌ها یک شبکه نظارتی عمیق را ارائه دادند که توجه مکانی^۴ را به ویژگی‌های استخراج شده‌ی چند مقیاسه اعمال می‌کرد. همچنین Liu و همکاران [۲۸] یک

۳ Decoder

۴ Spatial attention

۱ Early Fusion

۲ Late Fusion

شبکه‌ی تماماً پیچشی را بر پایه ترکیب ویژگی‌های چندمقیاسه ارائه کردند که توجه طیفی^۱ و توجه مکانی را به‌منظور استخراج ویژگی‌های زمینه‌ای^۲ اعمال می‌کرد. در تحقیقی دیگر Cz و همکاران [۲۹] یک روش ترکیب ویژگی به‌صورت چندوجهی را پیشنهاد دادند که ویژگی‌های چندمقیاسه را از طریق راه‌های مختلف با یکدیگر ترکیب می‌کند. هدف از این کار جریان بهتر اطلاعات در شبکه پیچشی است. در همین راستا Chen و همکاران [۲] یک روش ترکیب ویژگی بر پایه مدل‌های تبدیل شونده^۳ را ارائه کردند که ویژگی‌های معنادار بیشتری در شناسایی تغییرات ارائه می‌کرد. در مقاله‌ای دیگر Bandara و همکاران [۳۰] شبکه کدگذار^۴ را که از شبکه‌های مرسوم پیچشی است، با یک شبکه کدگذار سلسله مراتبی از جنس مدل‌های تبدیل شونده جایگزین کردند. هدف از این کار استخراج ویژگی‌های چندمقیاسه با دامنه تغییرات بیشتر بوده‌است. بنابراین Liu و همکاران [۳۱] در سال ۲۰۲۲، یک شبکه ترکیب ویژگی ارائه کردند که به‌صورت دوگانه از شبکه‌های پیچشی و مدل‌های تبدیل شونده برای ترکیب ویژگی‌های بصری چندمقیاسه بهره می‌برد. با همه‌ی این موارد، روش‌های ارائه‌شده یا به‌طور تمام و کمال از ویژگی‌های چندمقیاسه بهره نمی‌برند و یا از لحاظ محاسباتی بهینه نیستند [۲]. از طرفی کاهش پی‌درپی ابعاد نقشه‌های ویژگی^۵ در این شبکه‌ها، دلیل از بین رفتن اطلاعات مکانی در تعیین تغییرات مرز عوارض است که یک محدودیت رایج محسوب می‌شود [۳۲].

باهدف کاهش محدودیت‌های ذکرشده، در این تحقیق یک مدل یادگیری عمیق بر پایه شبکه‌های پیچشی ارائه‌شده است که با کاهش تعداد مراحل کوچک‌سازی نقشه‌های ویژگی از نظر حد تفکیک مکانی، به مقابله با محو اطلاعات در فرایند یادگیری می‌پردازد. این مدل یک شبکه پیچشی اصلاح‌شده بر پایه معماری ResNet [۳۳] را به‌منظور استخراج ویژگی‌های اولیه سطح پایین و سطح بالا از تصاویر دو زمانه بکار می‌گیرد. بر روی قدر مطلق تفاضل ویژگی‌های سطح بالا، چندین عملگر عمق‌مبنای، پیچشی با نرخ‌های رقیق‌سازی^۶ مختلف و عملگر نقطه‌ای اعمال می‌گردد تا علاوه بر کاهش مراحل کم کردن ابعاد نقشه‌های ویژگی، ویژگی‌های چند مقیاسی را به همراه

ویژگی‌های نقطه‌ای در مقدار محاسبات بهینه، استخراج کند. پس‌از آن ویژگی‌های سطح پایین با ویژگی‌های چند مقیاسی خروجی در مرحله قبل باهدف افزایش حد تفکیک ویژگی‌های تولیدی و جلوگیری از محوشدن عوارض تغییر کرده با ابعاد کوچک، ترکیب می‌گردد و پس از یک‌لایه عملگر پیچشی، نقشه تغییرات محاسبه می‌شود. استخراج ویژگی‌های چند مقیاسی با میزان محاسبات بهینه در کنار کاهش تعداد مراحل کم کردن ابعاد نقشه‌های ویژگی، باعث افزایش دقت و کاهش محاسبات موردنیاز جهت تولید نقشه تغییرات شده‌است.

نوآوری و اقدامات صورت گرفته در راستای بهبود نقشه تغییرات تصاویر سنجش‌ازدوری در ادامه تشریح می‌شود. در این تحقیق، یک شبکه پیچشی نوین ارائه گردیده‌است که با استخراج ویژگی‌های چند مقیاسی در کنار کاهش تعداد مراحل کم کردن ابعاد نقشه‌های ویژگی، علاوه بر کاهش میزان محاسبات موردنیاز فرایند، باعث افزایش دقت نقشه تغییرات در تصاویر طبیعی می‌شود. به‌علاوه مدل پیشنهادی شامل چندین نوع عملگر پیچشی است. از جمله عملگر رقیق‌شده جهت استخراج ویژگی‌های (۱) چندمقیاسه (۲) عملگر عمق‌مبنا جهت ترکیب و استخراج ویژگی‌های طیفی (۳) عملگر نقطه‌ای جهت افزایش دقت شناسایی مرز عوارض (۴) یک عملگر Pooling جهت بهبود جریان گرادیان آموزشی، شامل می‌شود. درنهایت نتایج ارزیابی ۷ مدل شناسایی تغییرات شناخته‌شده بر روی ۲ مجموعه داده مختلف نشان‌دهنده برتری شبکه پیشنهادی هم در معیارهای ارزیابی و هم در میزان پیچیدگی محاسبات است.

۲- مبانی نظری تحقیق

۲-۱- روش پیشنهادی

با توجه به اهمیت بالای شناسایی تغییرات در کاربردهای فتوگرامتری و سنجش‌ازدور و وجود مشکلات و محدودیت‌ها در تهیه یک نقشه دقیق و کارآمد، در این تحقیق یک فرایند

۴ Incoeder

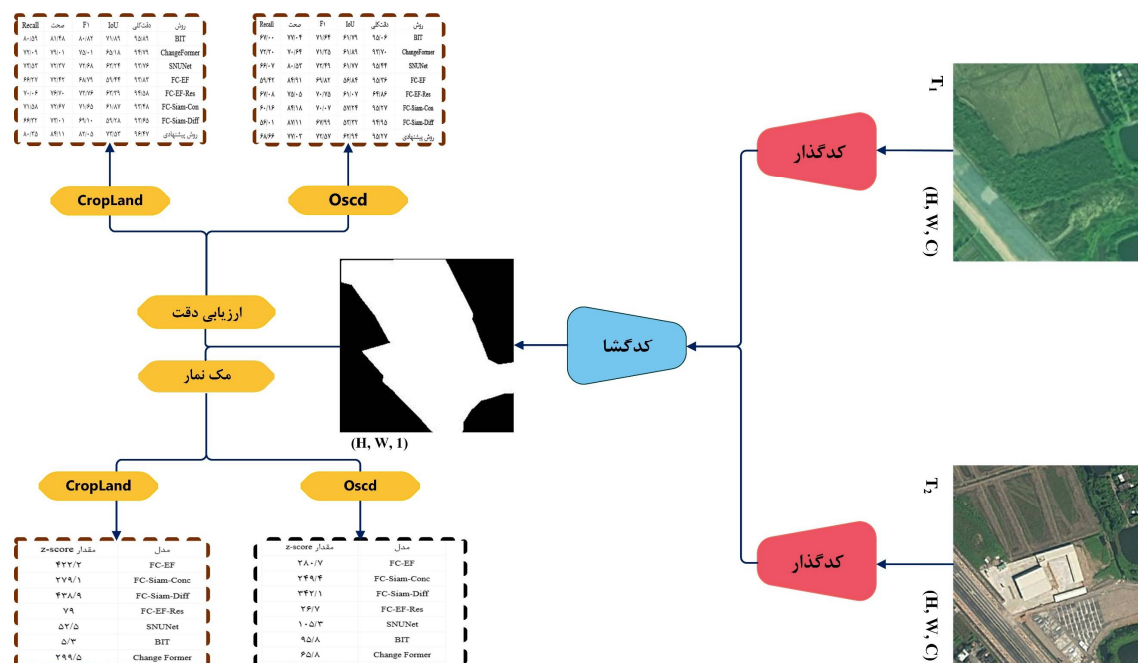
۵ Feature maps

۶ Dilution rate

۱ Channel attention

۲ Contextual features

۳ Transformers



شکل ۱- روندنمای کلی فرآیند طی شده در این تحقیق

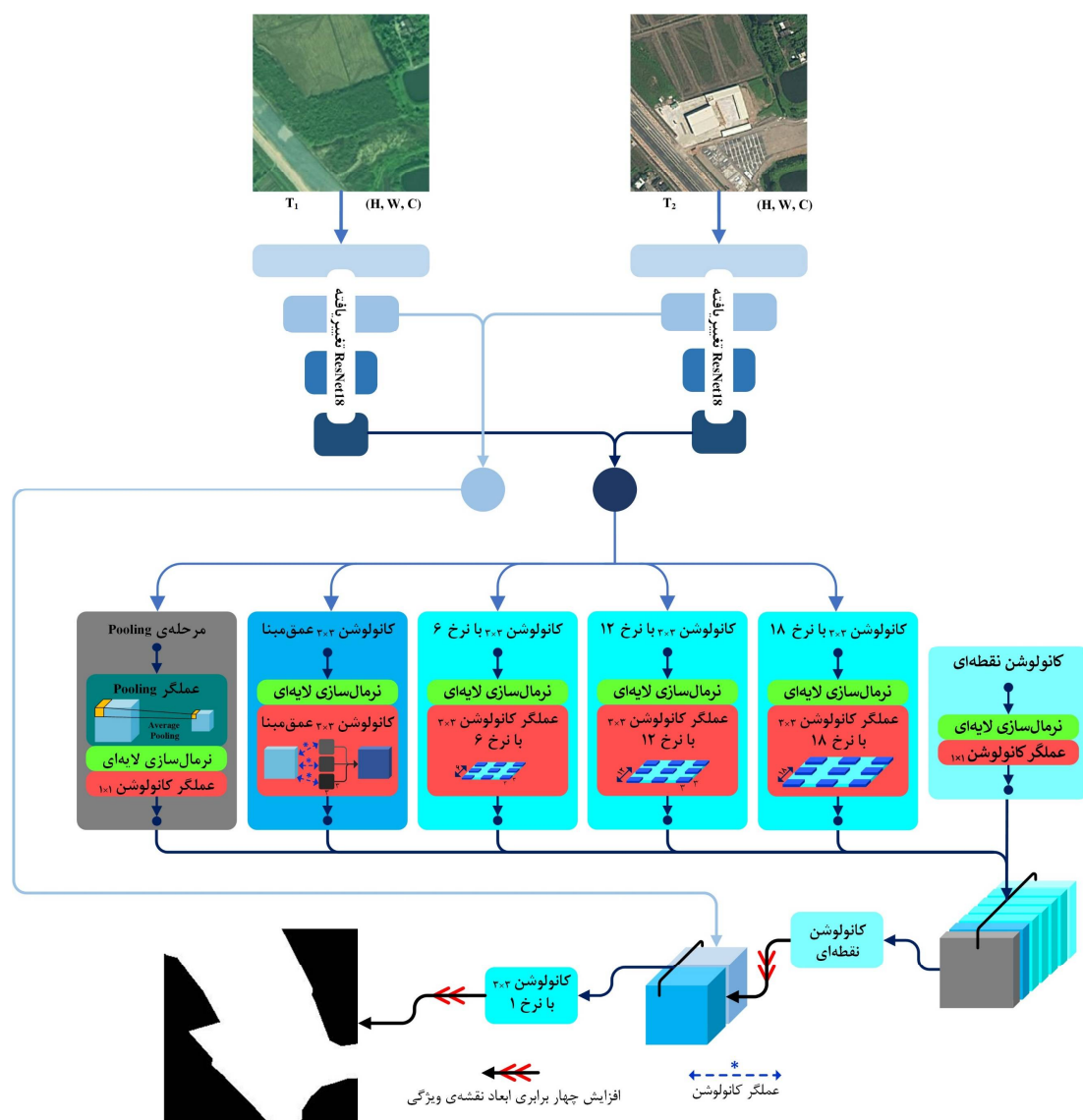
۲-۱-۱- استخراج ویژگی‌های دو زمانه

اولین قدم در تهیه نقشه تغییرات در روش‌های مبتنی بر شبکه‌های پیچشی، استخراج ویژگی است که معمولاً توسط یک شبکه پیچشی از پیش آموزش‌دیده صورت می‌پذیرد. این شبکه به‌طور هم‌زمان بر روی تصویر زمان اول و دوم اعمال می‌گردد و بخش مهمی از محاسبات مورد نیاز فرایند را به خود اختصاص می‌دهد. از این‌رو انتخاب و بهینه‌سازی مدل استخراج گر ویژگی یکی از اهداف محققان فعال در زمینه‌ی شناسایی تغییرات بوده‌است. در این تحقیق از معماری بهبود یافته ResNet18 استفاده شده‌است که در ابتدا یک عملگر پیچشی با ابعاد فیلتر 7×7 و طول گام^۱ دو بر روی ورودی اعمال می‌گردد. به دنبال آن یک نرمال‌ساز دسته‌ای^۲، تابع فعال‌سازی ReLU^۳ و عملگر Pooling قرار می‌گیرد که ابعاد نقشه‌های ویژگی استخراج‌شده را به یک‌چهارم ابعاد تصویر ورودی کاهش می‌دهد. در ادامه، ۴ بلوک باقیمانده^۴ ویژگی‌های سطح بالاتر را استخراج می‌کنند که هر کدام شامل دو عملگر پیچشی 3×3 ، دو نرمال‌ساز دسته‌ای و یک تابع فعال‌سازی ReLU می‌شود. در هر بلوک، نقشه ویژگی

بهینه طراحی شده‌است. روندنمای کلی این فرایند که در شکل (۱) قابل مشاهده است. مطابق روندنما، تصاویر دوزمانه‌ی مربوط به مجموعه داده CropLand [۳۱] در ابتدا از شبکه‌های کدگذار که یکی از دو بخش اصلی مدل‌های پس‌ترکیبی هستند، عبور می‌کنند و ویژگی‌های دوزمانه استخراج می‌گردند. سپس شبکه کدگشا، دیگر بخش اصلی، بر روی ویژگی‌های استخراج‌شده در مرحله قبل اعمال و بازسازی ویژگی انجام می‌شود. خروجی این قسمت نقشه تغییراتی خواهد بود که در مرحله آخر با داده‌های واقعیت زمینی از نظر کیفی و کمی مقایسه و ارزیابی خواهد شد. مدل ارائه‌شده در این تحقیق از دو بخش اصلی تشکیل شده است. که شامل قسمت استخراج ویژگی‌های دو زمانه و ترکیب و بازسازی ویژگی‌های شناسایی‌شده جهت محاسبه نقشه تغییرات می‌باشد. در این مدل در ابتدا ویژگی‌های اولیه توسط استخراج گرهای ویژگی یا کدگذارها به دست می‌آیند. سپس ویژگی‌های تولیدشده وارد فرایند بازسازی چندمقیاسی می‌شوند و پس از ترکیب با ویژگی‌های سطح پایین و اعمال یک عملگر پیچشی، خروجی نقشه تغییرات نهایی محاسبه می‌گردد. نمای کلی از فرایند مذکور و معماری طراحی‌شده در شکل (۲) قرار داده شده است.

^۳ Rectified Linear Unit
^۴ Residual Block

^۱ Stride
^۲ Batch Normalization



شکل ۲- روندنمای مدل پیچشی پیشنهادی

مرحله، نقشه‌های به‌دست‌آمده در چندین سطح پردازش‌شده نقشه تغییرات به دست می‌آید. در فرایند بازسازی طراحی‌شده، در ابتدا قدر مطلق تفاضل ویژگی‌های اولیه برای تصاویر زمان اول و دوم به‌صورت نقطه‌ای محاسبه می‌گردد. سپس نقشه‌های تفاضلی از بلوک باقیمانده سوم جهت استخراج و ترکیب ویژگی‌های طیفی و مختلف مقیاس، وارد زیر شبکه‌ی چندمقیاسی پیشنهادی می‌شوند. در این زیر شبکه یک عملگر پیچشی نقطه‌ای جهت افزایش توانایی مدل در تفکیک مرز عوارض بکار می‌رود. عملگرهای نقطه‌ای برخلاف نوع معمولی، از پنجره‌های همسایگی به‌عنوان مثال 3×3 جهت محاسبه خروجی استفاده نمی‌کنند و در نتیجه میزان ابهام در تفکیک مرزها کاهش می‌یابد. در

ورودی پس از عبور از یک سلسله عملگر پیچشی و نرمال‌سازی، با نقشه ورودی اولیه جمع می‌شوند و تابع فعال - ساز ReLU بر روی خروجی اعمال می‌گردد. با هدف کاهش تعداد لایه‌های کاهش‌دهنده ابعاد نقشه‌های ویژگی و در کنار حفظ بهینه بودن محاسبات، بلوک باقی‌مانده چهارم حذف شد و در اولین عملگر باقیمانده، طول گام دو با نرخ رقیق - سازی دو جایگزین گردید. در نهایت ویژگی‌های به‌دست‌آمده از بلوک اول و بلوک سوم، وارد فرایند بازسازی ویژگی خواهند شد.

۲-۱-۲- بازسازی ویژگی

مرحله بازسازی ویژگی، وظیفه غنی‌سازی و افزایش حد تفکیک ویژگی‌های تولیدشده در مرحله قبل را دارد. در این

ادامه، عملگر پیچشی عمق مبنا^۱ با ترکیب اطلاعات نسبت به بعد نقشه های ویژگی، بهره وری از ویژگی های چندمقیاسه و طیفی را افزایش می دهد [۳۴]. به علاوه در زیر شبکه ی مذکور، سه نوع عملگر پیچشی با نرخ های رقیق سازی ۶، ۱۲ و ۱۸ باهدف استخراج ویژگی های مختلف مقیاس و افزایش دقت مدل در مناطقی که تغییرات در ابعاد مختلف رخ می دهد، اعمال شده است. در قسمت آخر زیر شبکه ی پیشنهادی نیز یک لایه Pooling و عملگر پیچشی ۱×۱ قرار گرفته که باعث بهبود جریان گرادین آموزش و کاهش اثر محوشدگی آن می شود. نقشه های تولیدی از تمامی این عملگرها در کنار هم قرار گرفته و تعداد آن پس اعمال یک عملگر پیچشی نقطه ای به یک پنجم کاهش می یابد. خروجی این قسمت با نقشه های تفاضلی از بلوک باقیمانده اول ترکیب شده و با اعمال یک لایه پیچشی ۳×۳، خروجی نهایی نقشه ی تغییرات تشکیل می گردد.

۲-۲- روش های مقایسه ای

باهدف بررسی دقیق عملکرد مدل پیشنهادی، در این مقاله هفت مدل پیشرو^۲ در نظر گرفته شده است که در ادامه به شرح هر کدام پرداخته می شود.

۱. BIT [۲] یک مدل شناسایی تغییرات بر پایه تبدیل شونده ها است که با بهره وری کامل از مدل های کدگذار و کدگشا و نشانه های معنایی^۳ ویژگی های بصری کامل تری را در فضای ویژگی استخراج می کند. این روش باهدف ارائه مفاهیم سطح بالا، دسته ای از توکن ها را که تغییرات صورت گرفته در تصاویر دو زمانه را آشکار می کند، یاد می گیرد.

۲. مدل ChangeFormer [۳۰] از یک کدگذار سلسله مراتبی و دوگانه بر پایه مدل های تبدیل شونده برای استخراج ویژگی های دو زمانه استفاده کرده است. در قسمت کدگشا، یک مدل ساده ی چندلایه پرسپترون^۴ تعبیه شده است که ویژگی های چندمقیاسه را باهم ادغام می کند و شناسایی تغییرات را در سطوح مختلف انجام می دهد.

۳. مدل SNUNet [۳۵] از ترکیب یک شبکه ی تماماً

پیچشی Siamese، واحد برپایه ی مکانیسم توجه ECAM و یک شبکه NestedUNet برای استخراج ویژگی های چندمقیاسه و سطح بالا استفاده می کند. تابع هزینه دوگانه بکار رفته در این روش، مجموع توابع هزینه Cross Entropy وزن دار و تابع هزینه Dice می باشد.

۴. مدل FC-EF [۱۹] یک مدل تماماً پیچشی ای است که بر پایه ی معماری Unet [۳۶] است و ورودی آن ترکیب تصاویر دوزمانه از سمت کانال های تصویر است.

۵. مدل FC-EF-Res [۳۷] یک متغیر برای مدل FC-EF است که ایده بلوک های باقی مانده را به معماری Unet اضافه کرده و فرایند انتقال ویژگی های تغییر یافته را در حین آموزش بهبود بخشیده است.

۶. مدل FC-Siam-Conc [۱۹] حالت پس ترکیبی مدل FC-EF است و شناسایی تغییرات را با تفاضل ویژگی های ویژگی های دو زمانه چند سطحی از یک شبکه کاملاً پیچشی دوگانه، انجام می دهد.

۷. مدل FC-Siam-Diff [۱۹] حالت پس ترکیبی مدل FC-EF است و شناسایی تغییرات را با تفاضل ویژگی های دو زمانه چند سطحی از یک شبکه کاملاً پیچشی دوگانه، انجام می دهد.

مدل پیشنهادی و تمام آزمایش های مربوطه در چهارچوب PyTorch پیاده سازی شدند. اندازه دسته ای ۱۶ و نرخ یادگیری ۰/۰۰۱ برای تمام آموزش های مدل با استفاده از بهینه ساز Adam اتخاذ شده است. فرآیند آموزش ۵۰ دوره به طول می انجامد و استراتژی های تقویت داده^۵ از جمله چرخش عمودی و افقی^۶ به طور تصادفی به مجموعه داده ی آموزشی اعمال می شوند تا از تطبیق بیش از حد^۷ جلوگیری شود.

۲-۳- معیارهای ارزیابی دقت

چهار معیار رایج از جمله دقت^۸، Recall، معیار F۱، معیار دقت کلی^۹ و معیار نسبت تقاطع به اجتماع^{۱۰} برای ارزیابی دقت انتخاب شده اند که آن ها را می توان به صورت زیر تعریف کرد:

۷ Overfitting

۸ Precision

۹ Overall Accuracy

۱۰ Intersection or union

۱ Depthwise

۲ State of the art

۳ Semantic tokens

۴ Multi Layer Perceptron

۵ Data augmentation

۶ Horizontal and Vertical Flip

ارائه یک استاندارد برای مقایسه اثربخشی الگوریتم‌های مختلف شناسایی تغییرات می‌باشد و این مجموعه تمرکز خاصی بر روی تغییر مناطق شهری دارد. نمونه‌های موجود در این مجموعه، برای هر منطقه، ۲۶ باند را شامل می‌شوند که ۱۳ باند اول مربوط به تصویر زمان اول و ۱۳ باند باقی‌مانده مربوط به تصویر زمان دوم است. داده‌های واقعیت زمینی به صورت دستی تعیین گردیده و موقعیت مکانی و زمانی تمام نمونه‌ها مشخص می‌باشد.

در قسمت دوم این مجموعه، Ebel و همکاران [۳۸] در سال ۲۰۲۱ باهدف استفاده از خصوصیات چند منبعی در شناسایی تغییرات، داده ماهواره‌ی سنتینل-۱ را با توجه به مشخص بودن زمان و مکان نمونه‌های قسمت اول، به مجموعه‌ی OSCD اضافه کردند. تصاویر راداری سنتینل-۱ (مدار سعودی) از طریق موتور Google Earth Engine [۳۹] دانلود شده و ارتباط بین سیستم مختصات داده‌های نوری اصلی و مشاهدات راداری توسط برنامه GDAL [۴۰] صورت پذیرفته‌است.

$$PRE = \frac{TP}{TP + FP} \quad (۱)$$

$$REC = \frac{TP}{TP + FN} \quad (۲)$$

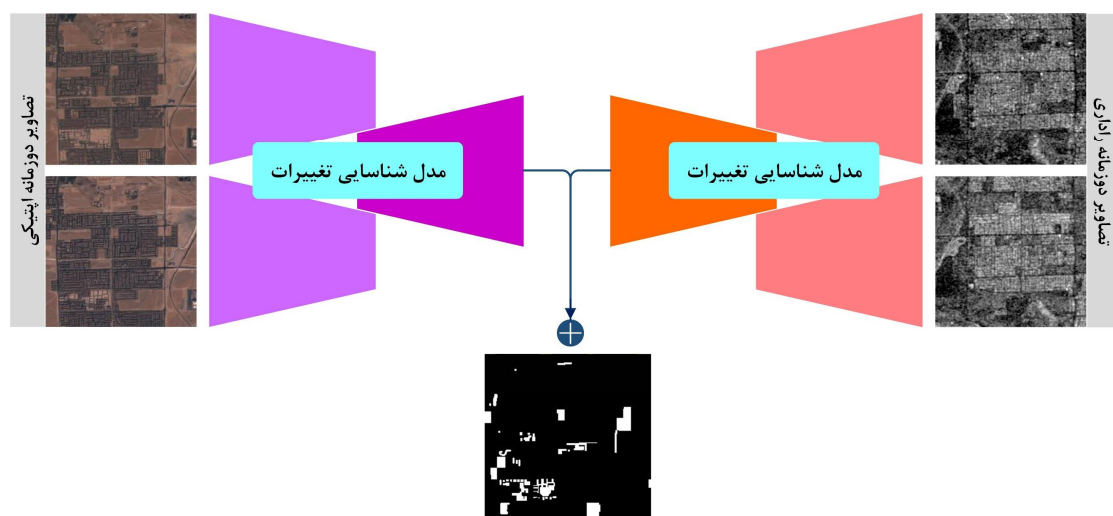
$$F1 = \frac{2 \times PRE \times REC}{PRE + REC} \quad (۳)$$

$$IOU = \frac{TP}{FP + TP + FN} \quad (۴)$$

$$Acc = \frac{TP + TN}{TP + TN + FN + FP} \quad (۵)$$

۲-۴- داده‌های مورد استفاده

۱. مجموعه OSCD [۹] به همراه تصاویر راداری: قسمت اول این مجموعه مربوط به تصاویر سنتینل-۲ است که در قالب مجموعه‌ی OSCD ارائه می‌شود. هدف از آن



شکل ۳- فرایند تهیه نقشه تغییرات در مجموعه OSCD

گرفته شده‌است. همچنین تصاویر موجود توسط ماهواره Gaofen۲ در استان گوانگدونگ کشور چین، در سال‌های ۲۰۱۷ و ۲۰۱۹ جمع‌آوری شدند که وضوح مکانی بین ۵/۵ تا ۲ متر را در برمی‌گیرد. لازم به ذکر است نمونه‌های موجود در این مجموعه شامل دو تصویر ۵۱۲×۵۱۲ پیکسل و یک داده واقعیت زمینی مربوط به تغییر زمین‌های زراعی می‌شوند.

همچنین در این مجموعه به دلیل چند منبعی بودن تصاویر، خروجی نهایی نقشه‌ی تغییرات از مجموع نتایج اعمال مدل بر روی تصاویر دو زمانه‌ی اپتیکی و راداری به دست می‌آید که روند نمای آن در شکل (۳) قرار گرفته‌است. ۲. مجموعه داده CropLand: این مجموعه شامل ۵۶۰ جفت تصویر دو زمانه از زمین‌های زراعی تغییر کرده می‌باشد که با نسبت ۱۲۰:۱۲۰:۳۲۰ نمونه به ترتیب برای قسمت‌های آموزش، اعتبارسنجی و ارزیابی مدل در نظر

۳- نتایج و بحث

در این بخش به بررسی روش های ارزیابی و مقایسه آن با سایر روش ها می پردازیم. با بررسی مدل های به کار برده شده در مجموعه OSCD، نتایج به دست آمده از مقایسه روش پیشنهادی با سایر روش های ارزیابی در جدول (۱) قرار گرفته است. این روش در معیارهای F1 و IoU به ترتیب مقادیر ۷۲/۵۷ و ۶۲/۹۴ درصد را به خود

اختصاص داده است که ۰/۰۸ و ۱/۱۷ درصد بالاتر از مدل SNUNet قرار گرفته در رتبه دوم معیار F1 است. به طور کلی پس از مدل پیشنهادی مدل های SNUNet و BIT به ترتیب با مقادیر ۷۲/۴۹ و ۷۱/۶۴ درصد برای معیار F1، بهتر از سایر مدل های ارزیابی عمل می کنند. همچنین کمترین دقت به دست آمده مربوط به مدل های FC-Siam-Diff و FC-EF است که به ترتیب دارای مقادیر F1، ۶۷/۹۹ و ۶۹/۸۲ درصد می باشند.

جدول ۱- نتایج به دست آمده بر روی مجموعه OSCD

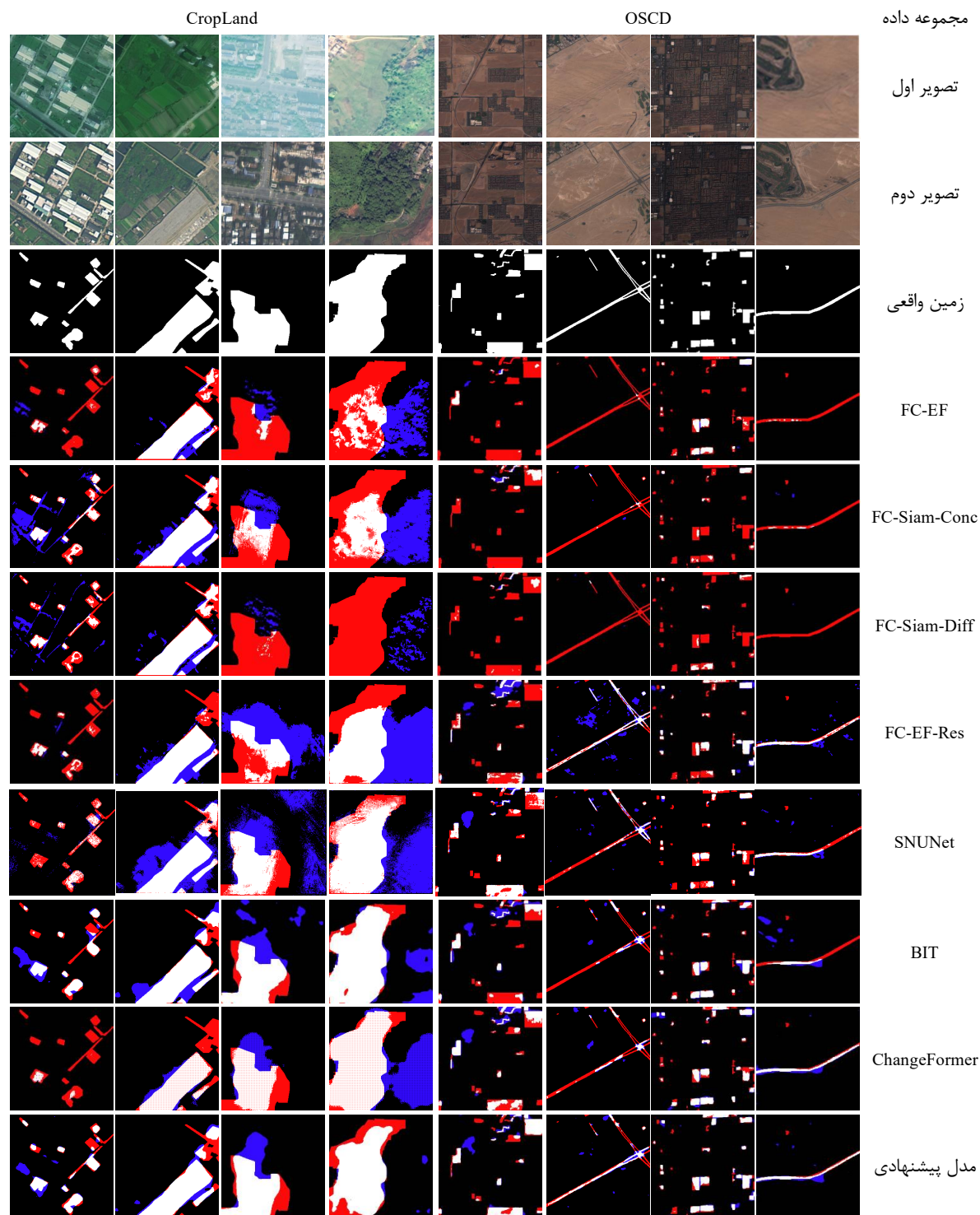
روش	دقت کلی	IoU	F1	دقت	Recall
BIT	۹۵/۰۶	۶۱/۷۹	۷۱/۶۴	۷۷/۰۴	۶۷/۰۰
ChangeFormer	۹۳/۷۰	۶۱/۸۹	۷۱/۳۵	۷۰/۶۴	۷۲/۲۰
SNUNet	۹۵/۴۴	۶۱/۷۷	۷۲/۴۹	۸۰/۵۳	۶۶/۰۷
FC-EF	۹۵/۳۶	۵۶/۸۴	۶۹/۸۲	۸۴/۹۱	۵۹/۴۲
FC-EF-Res	۶۴/۸۶	۶۱/۰۷	۷۰/۷۵	۷۵/۰۵	۶۷/۰۸
FC-Siam-Con	۹۵/۲۷	۵۷/۲۴	۷۰/۰۷	۸۴/۱۸	۶۰/۱۶
FC-Siam-Diff	۹۴/۹۵	۵۳/۳۲	۶۷/۹۹	۸۷/۱۱	۵۶/۰۱
روش پیشنهادی	۹۵/۲۷	۶۲/۹۴	۷۲/۵۷	۷۷/۰۳	۶۸/۶۶

جدول ۲- نتایج به دست آمده بر روی مجموعه CropLand

روش	دقت کلی	IoU	F1	دقت	Recall
BIT	۹۵/۸۹	۷۱/۸۹	۸۰/۸۲	۸۱/۴۸	۸۰/۵۹
ChangeFormer	۹۴/۷۹	۶۵/۱۸	۷۵/۰۱	۷۹/۰۱	۷۲/۰۹
SNUNet	۹۳/۷۶	۶۳/۲۴	۷۲/۶۸	۷۲/۳۷	۷۳/۵۳
FC-EF	۹۳/۸۳	۵۹/۴۴	۶۸/۷۹	۷۲/۴۲	۶۶/۲۷
FC-EF-Res	۹۴/۵۸	۶۳/۳۹	۷۲/۷۶	۷۶/۷۰	۷۰/۰۶
FC-Siam-Con	۹۳/۴۸	۶۱/۸۷	۷۱/۶۵	۷۲/۶۷	۷۱/۵۸
FC-Siam-Diff	۹۳/۶۵	۵۹/۲۸	۶۹/۱۰	۷۳/۰۱	۶۶/۳۲
روش پیشنهادی	۹۶/۴۷	۷۳/۵۳	۸۲/۰۵	۸۴/۱۱	۸۰/۳۵

برای مجموعه داده CropLand مدل پیشنهادی توانسته در معیار دقت کلی، IoU، F1 و صحت به ترتیب مقادیر ۹۶/۴۷، ۷۳/۵۳، ۸۳/۰۵ و ۸۱/۱۱ درصد را به دست آورد که F1 آن ها نسبت به مدل BIT و ChangeFormer به ترتیب حدود ۱/۲۳ و ۷/۰۴ درصد بیشتر است. همچنین کمترین F1 مربوط به مدل FC-EF با مقدار ۶۸/۷۹ درصد است که ۱۳/۲۶ درصد نسبت به مدل پیشنهادی ضعیف تر عمل می کند. خروجی های ظاهری مدل پیشنهادی و سایر

مدل های ارزیابی در شکل (۴) نشان داده شده است. با توجه به آن، برتری روش پیشنهادی در مناطقی که تغییرات مقیاس در عوارض تغییر کرده وجود دارد، دیده می شود. برای درک راحت تر، از سه رنگ مختلف برای تفسیر نتایج استفاده می شود: رنگ آبی برای نمایش مناطقی است که خروجی مدل، عوارض تغییر کرده را شناسایی می کند اما داده های واقعیت زمینی تغییری نشان



شکل ۴- خروجی‌های ظاهری برای روش پیشنهادی و روش‌های مقایسه‌ای

نمی‌دهند. رنگ سفید گویای محدوده‌ای می‌باشد که خروجی مدل و داده‌های واقعیت زمینی هر دو به عوارض تغییر کرده اشاره می‌کنند. همچنین، مناطقی که داده‌های واقعیت زمینی عوارض تغییر کرده را شناسایی کرده‌اند اما خروجی مدل تغییری نشان داده‌است، با رنگ قرمز مشخص شده‌اند. در ستون چهارم

مجموعه OSCD، جاده‌ای چهارراه مشاهده می‌شود که در مدل FC-Siam-Diff، FC-Siam-Conc، FC-EF بسیار پایینی (رنگ قرمز) دارد. اما در مدل پیشنهادی خروجی آن و داده‌های واقعیت زمینی هر دو عوارض تغییر کرده را به خوبی شناسایی کرده‌اند و میزان مناطق آبی و قرمز آن به حداقل رسیده‌است. همان‌طور که در ستون

پنجم نشان داده شده، روش پیشنهادی در مقایسه با SNUNet و ChangeFormer می تواند ناحیه تغییر را با دقت بیشتری شناسایی کند. نتایج تغییر توسط FC-Siam-Conc و FC-EF نسبتاً تکه تکه و دارای طبقه بندی نادرست شدید هستند که یادآور جدول (۱) و دقت پایین مدل های ذکر شده است.

خروجی تصاویر مجموعه CropLand در شکل (۴) نشان داده شده است. مشاهده می شود در این مجموعه، مدل پیشنهادی نسبت به FC-Siam-Diff، FC-Siam-Conc و FC-EF نواحی تغییر را که مربوط به فضای سبز (ستون اول) می باشد با دقت بالاتری شناسایی کرده است. همین طور دارای

مناطق آبی و قرمز کمتری است که نشان دهنده دقت بالاتر این مدل نسبت به سایر مدل های ارزیابی می باشد. همچنین این مجموعه نتایج بهتری هم برای شناسایی لبه های عوارض و هم برای کاهش تشخیص کاذب دارد. با توجه به شکل (۴) در این مجموعه، خروجی های مدل BIT از نظر دقت در شناسایی تغییرات نزدیکی بیشتری به مدل پیشنهادی دارد و می توان آن را در رتبه دوم قرارداد. به منظور درک عمیق تر مدل های شناسایی تغییرات مختلف در کاربردهای عملی، از دو معیار میزان عملیات بر روی اعداد اعشاری (FLOPs) و تعداد پارامترها (Params) استفاده می شود تا کارایی همه روش های مقایسه ای بیشتر مورد بررسی قرار گیرد.

جدول ۳- میزان پیچیدگی محاسبات و پارامترهای آموزشی مدل پیشنهادی و سایر مدل های ارزیابی

CropLand		OSCD		زمان اعمال مدل (ثانیه)	FLOPs (G)	پارامترهای آموزشی (M)	زمان آموزش مدل (ساعت)	مدل
IoU	F1	IoU	F1					
۵۹/۴۴	۶۸/۷۹	۵۶/۸۴	۶۹/۸۲	۰/۲۶۷	۷/۱۲۵	۱/۳۵۱	۰/۲۷	FC-EF
۶۱/۸۷	۷۱/۶۵	۵۷/۲۴	۷۰/۰۷	۰/۳۵۶	۱۰/۶۱۵	۱/۵۴۶	۰/۳۶	FC-Siam-Conc
۵۹/۲۸	۶۹/۱۰	۵۳/۳۲	۶۷/۹۹	۰/۳۴۷	۹/۴۰۷	۱/۳۵۱	۰/۳۳	FC-Siam-Diff
۶۳/۳۹	۷۲/۷۶	۶۱/۰۷	۷۰/۷۵	۰/۱۷۲	۹/۱۰۷	۱/۱۰۳	۰/۲۹	FC-EF-Res
۶۳/۲۴	۷۲/۶۸	۶۱/۷۷	۷۲/۴۹	۲/۱۵۴	۱۰۹/۴۹۹	۱۲/۰۳۴	۲/۱۸	SNUNet
۷۱/۸۹	۸۰/۸۲	۶۱/۷۹	۷۱/۶۴	۰/۶۴۴	۲۱/۲۰۸	۳/۴۹۵	۰/۸۴	BIT
۷۳/۵۳	۷۵/۰۱	۶۱/۸۹	۷۱/۳۵	۰/۶۸۴	۱۶/۷۹۸	۲۴/۲۹۶	۱/۰۸	Change Former
۷۳/۵۳	۸۲/۰۵	۶۲/۹۴	۷۲/۵۷	۰/۴۰۳	۱۹/۰۵۸۶	۹/۰۸۷	۰/۳۳	مدل پیشنهادی

جدول ۵- نتایج بدست آمده از آزمون مک نمار بر روی

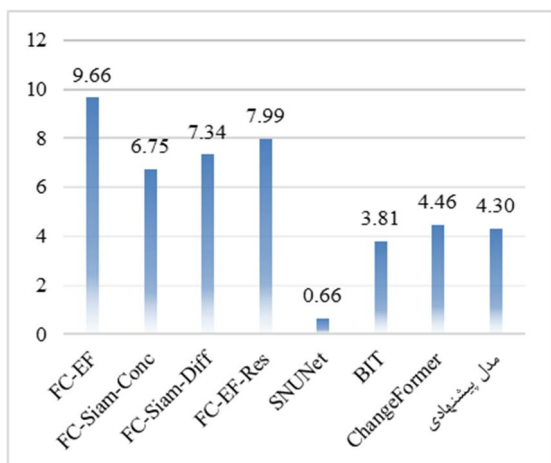
مجموعه CropLand

مقدار z-score	مدل
۲/۴۲۲	FC-EF
۱/۲۷۹	FC-Siam-Conc
۹/۴۳۸	FC-Siam-Diff
۷۹	FC-EF-Res
۵/۵۲	SNUNet
۳/۵	BIT
۵/۲۹۹	Change Former

جدول ۴- نتایج بدست آمده از آزمون مک نمار بر روی

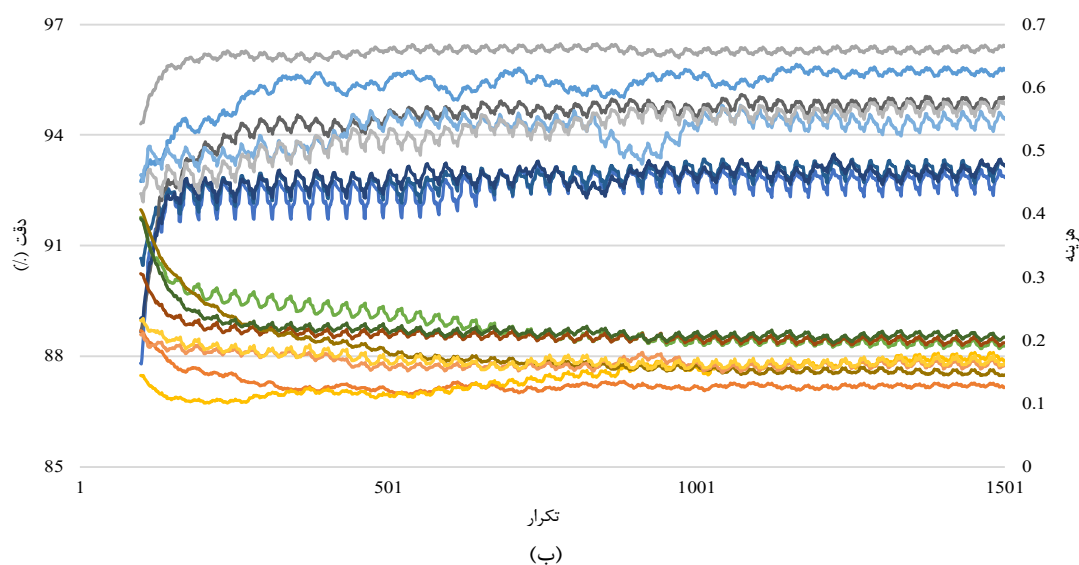
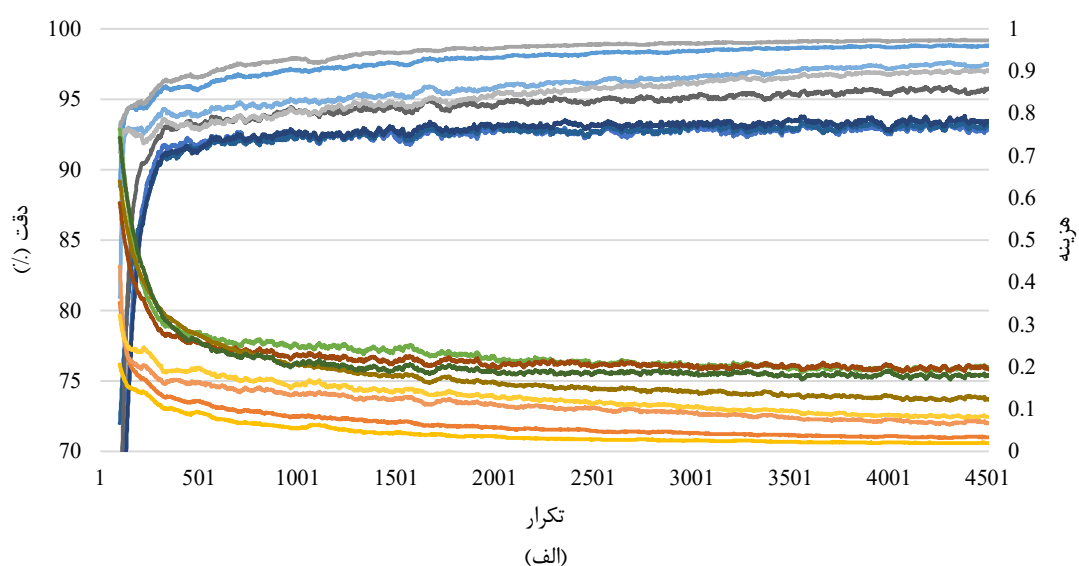
مجموعه OSCD

مقدار z-score	مدل
۲۸۰/۷	FC-EF
۲۴۹/۴	FC-Siam-Conc
۳۴۲/۱	FC-Siam-Diff
۲۶/۷	FC-EF-Res
۱۰۵/۳	SNUNet
۹۵/۸	BIT
۶۵/۸	Change Former



شکل ۵- نمودار F1/FLOPs برای روش پیشنهادی و مقایسه‌ای

پارامتر FLOPs پیچیدگی محاسباتی مدل را با محاسبه عملیات ضرب و جمع می‌سنجند که واحد آن میلیارد (G) می‌باشد. Params تعداد پارامترهایی است که باید در طول آموزش مدل، مطابق با پیچیدگی فضای آن، در واحد میلیون (M) آموزش داده شود. با توجه به دو ورودی دو زمانه با ابعاد $256 \times 256 \times 3$ پیکسل، میزان FLOPs و پارامترهای آموزشی همه روش‌ها در جدول (۳) آورده شده‌است که مشاهده می‌شود FC-EF، FC-EF-Res، BIT و SNUNet کمترین مقدار را دارند. با این حال، مدل پیشنهادی به دلیل استفاده از ترکیب ویژگی چند مقیاسی، بالاترین FLOPs را به خود اختصاص داده‌اند.



شکل ۶- نمودار مقادیر دقت و هزینه برای (الف) داده‌های آموزشی و (ب) داده‌های ارزیابی مجموعه CropLand

همچنین، بالاترین پارامترهای آموزشی مربوط به مدل ChangeFormer است. مدل پیشنهادی در کنار کسب بیشترین مقادیر در معیارهای ارزیابی، نسبت به برخی مدل های شناخته شده مانند BIT و ChangeFormer، به زمان کمتری جهت محاسبه خروجی های نقشه تغییرات نیاز دارد. جهت بررسی اثربخشی هر روش به ازای محاسبات مورد نیاز آن، می توان دقت F1 را به میزان عملیات های ریاضی مورد نیاز برای محاسبه ی خروجی، تقسیم کرد [۳۰، ۳۱]. هرچقدر مقدار این نسبت بالاتر باشد، اثربخشی آن نیز بیشتر است. در شکل (۵)، مقادیر به دست آمده برای این نسبت در مجموعه داده CropLand قابل مشاهده است. در این ارزیابی، SNUNet و FC-EF به ترتیب کم ترین و بیشترین مقدار را به خود اختصاص داده اند. روش پیشنهادی نیز با مقدار ۴/۳، بیشتر از BIT و کمتر از ChangeFormer است. در ادامه از آزمون مک نمار^۱ که شاید بهترین روش برای مقایسه دقیق نقشه موضوعی باشد [۴۱]، استفاده شده است. این آزمون یک آزمون آماری می باشد که برای بررسی معنی داری تست با استفاده از داده های زوج (دو روش مختلف در قیاس با داده واقعی زمینی) در طرح های مختلف مورد استفاده قرار می گیرد.

در حقیقت، آزمون مذکور از آزمون های آماری ناپارامتری است که برای ارزیابی همانندی دو نمونه استفاده می شود. بر اساس این تست می توان نتایج شناسایی تغییرات را با دو روش مختلف و به صورت دودویی باهم قیاس نمود و میزان سطح وابستگی آماری بین آن دو را مورد تحلیل قرار داد. بر این اساس، به منظور بررسی معنی داری اختلاف های ذکر شده در جدول ۴ و ۵، به بررسی موردی هر یک از روش های مورد قیاس با روش پیشنهادی در سطح معنی داری ۹۵٪ پرداخته شد.

در انتها نمودار دقت و هزینه تمامی مدل های مورد بررسی برای داده های آموزشی و ارزیابی مجموعه ی CropLand در شکل (۶) قرار گرفته است. همان طور که در این شکل قابل مشاهده است، برای داده های آموزشی، مدل پیشنهادی توانست دقت بالاتر و میزان خطای کمتری را به ثبت برساند. همین مورد برای داده های ارزیابی نیز رخ داده که می توان دلیل آن را در استفاده از عملگرهای رقیق شده، نقطه ای، عمقی و Pooling جهت بهبود جریان گرادین آموزش

جست و جو کرد. با این حال مدل FC-EF کمترین دقت و بیشترین مقدار هزینه را بر روی داده های آموزشی و ارزیابی از خود به جای گذاشته است.

۴- نتیجه گیری

شناسایی تغییرات یکی از اساسی ترین وظایف رشته فتوگرامتری و سنجش از دور می باشد که در زمینه های مختلف از جمله برنامه ریزی های شهری و کشوری از اهمیت بالایی برخوردار است. محققان در سال های اخیر با طراحی روش های مختلف سعی در افزایش دقت و کاهش میزان محاسبات مورد نیاز جهت محاسبه نقشه خروجی داشته اند. در همین راستا، در این تحقیق، یک مدل نوین شناسایی تغییرات بر پایه شبکه های عصبی پیچشی ارائه شده است که با استخراج و ترکیب ویژگی های چندمقیاسه و سلسله مراتبی باعث بهبود نتایج نهایی هم از نظر دقت و هم از نظر حجم و پیچیدگی محاسبات می شود. روش پیشنهادی چندین عملگر پیچشی را به طور هم زمان بر روی ویژگی های تولید شده در قسمت مدل کدگذار، اعمال می کند. از جمله این عملگرها می توان به عملگر رقیق شده جهت استخراج ویژگی های چندمقیاسه، عملگر عمق مبنا جهت ترکیب ویژگی های طیفی، عملگر نقطه ای جهت افزایش دقت شناسایی مرز عوارض و یک عملگر Pooling جهت بهبود جریان گرادین آموزش اشاره نمود.

جهت ارزیابی مدل ارائه شده، هفت مدل شناسایی تغییرات پیشرو و دو مجموعه داده به ترتیب با حد تفکیک مکانی متوسط و بسیار بالا در نظر گرفته شد. مقادیر به دست آمده برای معیار F1 در مجموعه اول برابر ۷۲/۵۷ و در مجموعه دوم ۸۲/۰۵ درصد می باشد که به ترتیب حاکی از بهبود ۰/۰۸ و ۱/۲۳ درصد نسبت به مدل قرار گرفته در جایگاه دوم معیار F1 است. بهره وری مناسب از ویژگی های چندمقیاسه، در کنار میزان محاسبات بهینه باعث برتری مدل پیشنهادی نسبت به سایر مدل های ارزیابی از نقطه نظر کیفیت ظاهری و نتایج عددی شده است. به عنوان پیشنهاد در تحقیقات آتی می توان از مدل های تبدیل شونده در بخش کدگذار و کدگشا استفاده و ویژگی های سطح بالاتری استخراج کرد. همچنین در این تحقیق فقط مدل های کدگذار پیش آموزش هستند. می توان با از پیش آموزش دادن وزن های مدل کدگشا دقت و کارایی مدل را افزایش داد.

^۱ McNemar

- [۱] C. Zhang et al., "A deeply supervised image fusion network for change detection in high resolution bi-temporal remote sensing images," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 166, pp. 183-200, 2020/08/01/ 2020, doi: <https://doi.org/10.1016/j.isprsjprs.2020.06.003>.
- [۲] H. Chen, Z. Qi, and Z. Shi, "Remote sensing image change detection with transformers," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-14, 2021.
- [۳] B. DeVries, M. Decuyper, J. Verbesselt, A. Zeileis, M. Herold, and S. Joseph, "Tracking disturbance-regrowth dynamics in tropical forests using structural change detection and Landsat time series," *Remote Sensing of Environment*, vol. 169, pp. 320-334, 2015.
- [۴] S. Wang, D. Quan, X. Liang, M. Ning, Y. Guo, and L. Jiao, "A deep learning framework for remote sensing image registration," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 145, pp. 148-164, 2018.
- [۵] M. Zhang and W. Shi, "A feature difference convolutional neural network-based change detection method," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 10, pp. 7232-7246, 2020.
- [۶] K. Nemoto, R. Hamaguchi, M. Sato, A. Fujita, T. Imaizumi, and S. Hikosaka, "Building change detection via a combination of CNNs using only RGB aerial imageries," in *Remote Sensing Technologies and Applications in Urban Environments II*, 2017, vol. 10431: SPIE, pp. 107-118.
- [۷] S. Ji, Y. Shen, M. Lu, and Y. Zhang, "Building instance change detection from large-scale aerial images using convolutional neural networks and simulated samples," *Remote Sensing*, vol. 11, no. 11, p. 1343, 2019.
- [۸] R. Liu, M. Kuffer, and C. Persello, "The temporal dynamics of slums employing a CNN-based change detection approach," *Remote Sensing*, vol. 11, no. 23, p. 2844, 2019.
- [۹] R. C. Daudt, B. Le Saux, A. Boulch, and Y. Gousseau, "Urban change detection for multispectral earth observation using convolutional neural networks," in *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, 2018: Ieee, pp. 2.۲۱۱۸-۲۱۲۰.
- [۱۰] F. Rahman, B. Vasu, J. Van Cor, J. Kerekes, and A. Savakis, "Siamese network with multi-level features for patch-based change detection in satellite imagery," in *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2018, IEEE, pp. 958-962.
- [۱۱] M. Wang, K. Tan, X. Jia, X. Wang, and Y. Chen, "A deep siamese network with hybrid convolutional feature extraction module for change detection based on multi-sensor remote sensing images," *Remote Sensing*, vol. 12, no. 2, p. ۲۰۲۰, ۲۰۲۰.
- [۱۲] H. Chen and Z. Shi, "A spatial-temporal attention-based method and a new dataset for remote sensing image change detection," *Remote Sensing*, vol. 12, no. 10, p. 1662, 2020.
- [۱۳] P. P. De Bem, O. A. de Carvalho Junior, R. Fontes Guimarães, and R. A. Trancoso Gomes, "Change detection of deforestation in the Brazilian Amazon using landsat data and convolutional neural networks," *Remote Sensing*, vol. 12, no. 6, p. 901, 2020.
- [۱۴] J. Chen et al., "DASNet: Dual attentive fully convolutional Siamese networks for change detection in high-resolution satellite images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 1194-1206, 2020.
- [۱۵] M. Zhang, G. Xu, K. Chen, M. Yan, and X. Sun, "Triplet-based semantic relation learning for aerial remote sensing image change detection," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 2, pp. 266-270, 2018.
- [۱۶] X. Peng, R. Zhong, Z. Li, and Q. Li, "Optical remote sensing image change detection based on attention mechanism and image difference," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 9, pp. 7296-7307, 2020.
- [۱۷] H. Jiang, X. Hu, K. Li, J. Zhang, J. Gong, and M. Zhang, "PGA-SiamNet: Pyramid feature-based attention-guided Siamese network for remote sensing orthoimagery building change detection," *Remote Sensing*, vol. 12, no. 3, p. 484, 2020.
- [۱۸] F. I. Diakogiannis, F. Waldner, and P. Caccetta, "Looking for change? Roll the dice and demand attention," *Remote Sensing*, vol. 13, no. 18, p. 3707, 2021.
- [۱۹] R. C. Daudt, B. Le Saux, and A. Boulch, "Fully convolutional siamese networks for change detection," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, 2018: IEEE, pp. 4063-4067.
- [۲۰] M. Lebedev, Y. V. Vizilter, O. Vygolov, V. Knyaz, and A. Y. Rubis, "CHANGE DETECTION IN REMOTE SENSING IMAGES USING CONDITIONAL ADVERSARIAL NETWORKS," *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, vol. 42, no. 2, 2018.
- [۲۱] D. Peng, Y. Zhang, and H. Guan, "End-to-end change detection for high resolution satellite images using improved UNet++," *Remote Sensing*, vol. 11, no. 11, p. 1382, 2019.

- [۲۲] T. Bao, C. Fu, T. Fang, and H. Huo, "PPCNET: A combined patch-level and pixel-level end-to-end deep network for high-resolution remote sensing image change detection," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 10, pp. 1797-1801, 2020.
- [۲۳] B. Hou, Q. Liu, H. Wang, and Y. Wang, "From W-Net to CDGAN: Bitemporal change detection via deep learning techniques," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 3, pp. 1790-1802, 2019.
- [۲۴] Y. Zhan, K. Fu, M. Yan, X. Sun, H. Wang, and X. Qiu, "Change detection based on deep siamese convolutional network for optical aerial images," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 10, pp. 1845-1849, 2017.
- [۲۵] B. Fang, L. Pan, and R. Kou, "Dual learning-based siamese framework for change detection using bi-temporal VHR optical remote sensing images," *Remote Sensing*, vol. 11, no. 11, p. 1292, 2019.
- [۲۶] P. F. Alcantarilla, S. Stent, G. Ros, R. Arroyo, and R. Gherardi, "Street-view change detection with deconvolutional networks," *Autonomous Robots*, vol. 42, no. 7, pp. 1301-1322, 2018.
- [۲۷] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested u-net architecture for medical image segmentation," in *Deep learning in medical image analysis and multimodal learning for clinical decision support*: Springer, 2018, pp. 3-11.
- [۲۸] Y. Liu, C. Pang, Z. Zhan, X. Zhang, and X. Yang, "Building change detection for remote sensing images using a dual-task constrained deep siamese convolutional network model," *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 5, pp. 811-815, 2020.
- [۲۹] J. Xu, C. Luo, X. Chen, S. Wei, and Y. Luo, "Remote sensing change detection based on multidirectional adaptive feature fusion and perceptual similarity," *Remote Sensing*, vol. 13, no. 15, p. 3053, 2021.
- [۳۰] W. G. C. Bandara and V. M. Patel, "A transformer-based siamese network for change detection," *arXiv preprint arXiv:2201.01293*, 2022.
- [۳۱] M. Liu, Z. Chai, H. Deng, and R. Liu, "A CNN-transformer Network with Multi-scale Context Aggregation for Fine-grained Cropland Change Detection," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2022.
- [۳۲] R. Shao, C. Du, H. Chen, and J. Li, "SUNet: Change Detection for Heterogeneous Remote Sensing Images from Satellite and UAV Using a Dual-Channel Fully Convolution Network," *Remote Sensing*, vol. 13, no. 18, p. 3750, 2021.
- [۳۳] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.
- [۳۴] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A convnet for the 2020s," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 11976-11986.
- [۳۵] S. Fang, K. Li, J. Shao, and Z. Li, "SNUNet-CD: A densely connected Siamese network for change detection of VHR images," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2021.
- [۳۶] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, 2015: Springer, pp. 234-241.
- [۳۷] R. C. Daudt, B. Le Saux, A. Boulch, and Y. Gousseau, "Multitask learning for large-scale semantic change detection," *Computer Vision and Image Understanding*, vol. 187, p. 102783, 2019.
- [۳۸] P. Ebel, S. Saha, and X. X. Zhu, "Fusing multi-modal data for supervised change detection," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 43, pp. 243-249, 2021.
- [۳۹] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, "Google Earth Engine: Planetary-scale geospatial analysis for everyone," *Remote sensing of Environment*, vol. 202, pp. 18-27, 2017.
- [۴۰] F. Warmerdam, "The geospatial data abstraction library," in *Open-source approaches in spatial data handling*: Springer, 2008, pp. 87-104.
- [۴۱] A. Kiani, F. Farnood Ahmadi, and H. Ebadi, "Correction of training process in object-based image interpretation via knowledge-based system capabilities," *Multimedia Tools and Applications*, vol. 80, no. 16, pp. 24901-24924, 2021/04/12 2021, doi: 10.1007/s11042-021-10824-0.