نشریه علمی

علوم و مهندسی آبخیزداری ایران
Iran-Watershed Management Science & Engineering

سال هجدهم- شماره ۶۷- زمستان ۱۴۰۳

Vol. 18, No. 67, Winter 2025

طبقه‌بندی سری زمانی کاربری اراضی با استفاده از شاخص‌های طیفی، تصاویر سنتینل ۲ و نمونه‌های آموزشی متغیر در بستر گوگل ارت انجین (GEE)

محمد کاظمی[۱*]، عاطفه جعفرپور[۲]

تاریخ دریافت: ۱۴۰۳/۰۵/۰۳ تاریخ پذیرش: ۱۴۰۳/۰۶/۱۳

https://doi.org/10.22034/18.67.1

چکیده مبسوط

مقدمه

سنجش از راه دور به ابزاری ارزشمند جهت تهیه داده‌های یکپارچه مکانی در مورد پوشش و کاربری زمین در مقیاس‌های زمانی و مکانی تبدیل شده است. یکی از چالش‌های اصلی در نقشه‌برداری چند-زمانه پوشش و کاربری زمین، در دسترس بودن و یکپارچگی داده‌های آموزشی برای الگوریتم‌های طبقه‌بندی نظارت شده است. جمع‌آوری نمونه‌های آموزشی برای هر کلاس پوشش و کاربری زمین در طول دوره‌های زمانی مختلف می‌تواند زمان بر و از لحاظ میدانی به‌ویژه در محیط‌های در حال تغییر سریع، چالش برانگیز باشد. این مسئله همچنین با امکان تغییرات طیفی و فنولوژیکی در ویژگی‌های پوشش زمین در طول زمان تشدید می‌شود و از این حیث می‌تواند قابلیت انتقال نمونه‌های آموزشی را کاهش دهد. مفهوم «مهاجرت یا انتقال» نمونه‌های آموزشی از یک سال مرجع به سال‌های هدف (مقصد) به عنوان راهی برای غلبه بر محدودیت داده‌های آموزشی مورد بررسی قرار گرفته است. در این میان، استفاده از سامانه گوگل ارت انجین (GEE) نقشه‌برداری چندزمانه پوشش و کاربری زمین را تسهیل کرده است. توانایی GEE در ادغام منابع داده مختلف، از جمله تصاویر Sentinel-2 و طیف گسترده‌ای از شاخص‌های طیفی، به محققان امکان توسعه کاربست طبقه‌بندی پوشش و کاربری زمین قوی و مقیاس‌پذیر را داده است. رصد تغییرات در استفاده و پوشش زمین در طول زمان برای درک و مدیریت محیط زیست حیاتی است. با این حال، هنگامی که محدودیت داده‌های آموزشی برای دوره‌های زمانی مختلف وجود داشته باشد، می‌تواند چالش‌برانگیز باشد. پژوهش حاضر رویکردی نوآورانه برای طبقه‌بندی تصاویر ماهواره‌ای Sentinel-2 در سال‌های مختلف با استفاده از مجموعه نمونه‌های آموزشی مرجع است.

مواد و روش‌ها

در پژوهش حاضر، کاربست نوآورانه‌ای با استفاده از نمونه‌های آموزشی متغیر از یک سال مرجع (تصاویر Sentinal-2 سال ۲۰۱۹)، همراه با باندهای تصاویر Sentinel-2 و شاخص‌های طیفی، برای طبقه‌بندی پوشش و کاربری زمین در منطقه مانگرو خوران به عنوان منطقه‌ای پویا و مهم اکولوژیکی مورد بررسی قرار گرفت. در این راستا، کاربرد الگوریتم‌های یادگیری ماشین در چارچوب سامانه GEE برای دستیابی به دقت طبقه‌بندی بالا و پایش تغییرات پوشش زمین در طول زمان بررسی شد. در این پژوهش، تصاویر ماهواره‌ای Sentinel-2 پوشش‌دهنده منطقه مورد مطالعه برای سال‌های هدف ۲۰۲۲ و ۲۰۲٤ و سال مرجع ۲۰۱۹ در GEE فراخوانی شدند. سپس جمع‌آوری داده‌های حقیقت زمینی شامل موقعیت و طبقه‌بندی انواع مختلف پوشش زمین برای سال مرجع با استفاده از نقشه کاربری اراضی سازمان فضایی اروپا جمع‌آوری شد. در ادامه، داده‌های حقیقت زمینی با کیفیت بالا و نمونه‌های تصویری متناظر آن‌ها از سال مرجع (۲۰۱۹) به عنوان نمونه‌های آموزشی به تصاویر سال هدف با الگوریتم فاصله زاویه طیفی (SAD) منتقل شدند. همچنین الگوریتم‌های طبقه‌بندی RF، GBT و CART برای طبقه‌بندی تصاویر سال هدف با استفاده از نمونه‌های آموزشی متغیر به کار گرفته شدند. تصاویر طبقه‌بندی شده با استفاده از معیارهای مختلف صحت، مانند دقت کلی، ضریب کاپا، دقت تولیدکننده و دقت کاربر ارزیابی و خطای گماشته شده و حذف شده ارزیابی شدند. در نهایت اهمیت باندهای طیفی مختلف Sentinel-2 و شاخص‌های طیفی در فرآیند طبقه‌بندی تجزیه و تحلیل شد تا مناسب‌ترین ویژگی‌ها برای تفکیک پدیده‌های مختلف منطقه مورد مطالعه شناسایی شوند.

نتایج و بحث

نتایج پژوهش حاضر نشان داد از میان الگوریتم‌های طبقه‌بندی، بیشترین میزان دقت برای صحت کلّی و کاپا تصاویر طبقه‌بندی شده سال ۲۰۲٤ و ۲۰۲۲، مربوط به طبقه‌بندی جنگل تصادفی به ترتیب با دقت ۰/۹۱۰٤ و ۰/۸۷٤۲، ۰/۸۹۵۵ و ۰/۸۵۷۰ بوده است. همچنین نتایج نشان داد مساحت جنگل‌های حرا در طول دوره بررسی از ۷۵۳۰/۷٤ هکتار به ٦٥٤٦/٥١ تنزل داشته است، چیزی در حدود ۹۸٤/۲۳ هکتار کاهش و به عبارتی ١٦٤ هکتار کاهش سطح جنگل‌های حرا به ازای هر

۱- نویسنده مسئول و دانشیار مرکز مطالعات و تحقیقات هرمز، دانشگاه هرمزگان، ایران mohamad.kazemi86@gmail.com

۲- استادیار، منابع طبیعی و محیط زیست، دانشگاه شهید باهنر کرمان، کرمان، ایران، Atefeh.jafarpoor@yahoo.com

کاهش و به عبارتی ۱٦٤ هکتار کاهش سطح جنگل‌های حرا به ازای هر سال. مساحت پهنه‌های مسکونی و عرصه‌های ساخت و ساز از ۷۲/۵۲۵۲ هکتار در سال ۲۰۱۹ به ۹٦/٤۸۱۵ در سال ۲۰۲٤ رسیده است که رشد شتابان این کاربری در دو سال آخر دیده می‌شود. نتیجه اشتراکات و جمع‌بندی روش‌های مختلف طبقه‌بندی در سنوات مختلف و اهمیت‌نسبی باندها و شاخص‌ها نشان داد که شاخص‌های EMVI و mNDWI بدلیل فراگیری پدیده‌های جنگل‌های مانگرو و نیز پهنه‌های آبی در محدوده مورد مطالعه غالبیت بیشتری از خود نشان داده‌اند. لذا در ترکیبات باندی جهت تفکیک پدیده‌های مختلف استفاده از باندهای سبز، قرمز، مادون قرمز نزدیک و از میان شاخص‌ها، شاخص مانگرو و نیز شاخص بهینه پهنه‌های آبی برای محدوده مطالعاتی مناسب‌ترین تشخیص داده شدند و برای عرصه‌های مشابه در جنوب ایران و محدوده‌های مانگرو پیشنهاد می‌شوند. در پژوهش حاضر با انتقال نقاط حقیقت زمینی سال منبع به تصاویر متناظر آن‌ها به عنوان نمونه‌های آموزشی در تصاویر سال هدف، از یک کاربست ساده و مؤثر برای طبقه‌بندی تصاویر ماهواره‌ای در پلتفرم GEE استفاده شد که قابلیت تسری به سایر مناطق را دارد.

نتیجه‌گیری

به طور کلی، پژوهش حاضر پتانسیل استفاده از نمونه‌های آموزشی متغیر و الگوریتم‌های پیشرفته یادگیری ماشین مانند جنگل تصادفی، درخت گرادیان تقویت شده و درخت رگرسیون و طبقه‌بندی و کاربرد شاخص‌های طیفی به عنوان داده‌های کمکی را برای طبقه‌بندی دقیق تصاویر ماهواره‌ای چند-زمانه نشان می‌دهد. توسعه ابزارهای مکانی، از جمله پلتفرم آنلاین GEE ، برای مدیریت به روز کاربری‌های زمین، از جمله مناطق تالابی و مانگرو، ضروری است. در پژوهش حاضر، نمونه‌های آموزشی با کیفیت بالا از سال مرجع به سال هدف (مهاجرت نمونه‌های آموزشی از زمان مشخص به سایر زمان‌ها) منتقل شدند و در انتها دقت طبقه‌بندی بالایی با استفاده از الگوریتم طبقه‌بندی جنگل تصادفی نسبت به سایر روش‌ها مانند رگرسیون تقویت‌شده یا درختان رگرسیون و طبقه‌بندی به دست آمد. این روش راه حل بالقوه‌ای در مطالعات چند-زمانه کاربری زمین و کمبود یا ناکافی بودن نمونه‌های آموزشی در سیستم GEE است. در این راستا پیشنهاد می‌شود که در مطالعات آینده، از ترکیبی از روش‌های فاصله اقلیدسی (ED)، فاصله زاویه طیفی (SAD) و خوشه‌بندی K-means برای تولید نمونه‌های آموزشی متغیر استفاده شود و نتایج طبقه‌بندی تصویر با استفاده از این روش‌ها مقایسه و تحلیل شود. این رویکرد راه‌حل امیدوارکننده‌ای برای تولید نقشه‌های روزآمد کاربری/ پوشش زمین، حتی در محیط‌های چالش برانگیز با داده‌های آموزشی محدود، ارائه می‌دهد. نتایج به دست آمده از این مطالعه می‌تواند جهت‌دهی مطالعات آتی را برای پایش پوشش اراضی، مدیریت و حفاظت مؤثر از منابع طبیعی ارزشمند، مانند جنگل‌های مانگرو را در سایر مناطق، هدایت کند.

کلیدواژه‌ها: ***الگوریتم‌های طبقه‌بندی پیکسل‌پایه، اهمیت نسبی، فاصله زاویه طیفی، کاربری اراضی، مانگرو، نمونه‌های آموزشی متغیر***

مقدمه

به‌هنگام‌رسانی نقشه‌های موضوعی از جمله نقشه کاربری‌های اراضی برای مدیریت منابع طبیعی و انسانی یکی از چالش‌های حوزه فتوگرامتری و سنجش از دور است. در این راستا، تحلیل تصاویر چندزمانه، امکان شناسایی و بررسی پدیده‌های پویا در سطح زمین را فراهم می‌کند [۲۸]. پردازش مجموعه داده‌های بزرگ سنجش از دور که طی سنوات و دهه‌های قبل جمع‌آوری شدند، به ابزارهای پیشرفته و قدرتمندی نیاز دارد [۱] که این موضوع می‌تواند با پلتفرم‌های رایانش ابری مانند گوگل ارت انجین[1] (GEE) حل شود [۱۱]. سامانه گوگل ارت انجین با ایجاد بستری مناسب به‌منظور پردازش سریع و ابری پروژه‌های سنجش از دور نقش مؤثری در زمینه تهیه نقشه‌های کارآمد و به‌روز ایفاد می‌کند [۲۵]. با این حال، کمبود نمونه‌های آموزشی ثابت کافی در زمان‌های مختلف یکی از محدودیت‌های قابل‌توجه برای پایش‌های سری‌زمانی تصاویر ماهواره‌ای و تهیه نقشه‌های موضوعی است [۸]. گوگل ارت انجین[2] (GEE) یک سامانه رایانش ابری برخط است که به کاربران اجازه دسترسی رایگان به داده‌های ماهواره‌ای بایگانی شده و بعضاً به‌روز را می‌دهد و به همراه ابزارهای متعدد پردازش درونی مانند الگوریتم‌های طبقه‌بندی امکان تحلیل مجموعه داده‌های بزرگ حاصل از تصاویر را فراهم می‌کند [۱۱]. الگوریتم‌های یادگیری ماشینی نظارت شده به دلیل عملکرد خوب و قدرتمند آن‌ها انتخاب‌های رایجی در طبقه‌بندی کاربری‌های اراضی محدوده‌های تالابی هستند [۳۳]، صرفنظر از مزایای این روش‌ها، نتایج نهایی تا حد زیادی به تعداد و کیفیت نمونه‌های آموزشی جهت آموختن یادگیری ماشینی به فرایند طبقه‌بندی دارد. بنابراین کمبود نمونه‌های تعلیمی با کیفیت بالا یک مشکل اساسی در مطالعاتی است که از این الگوریتم‌ها استفاده می‌کنند [۱۲]. این مشکل خصوصاً در مناطقی که نمونه‌های آموزشی قبلی (تاریخی) وجود ندارد دوچندان می‌شود [۲۳]. مواجه با این مشکل نیازمند روش‌های مناسبی است که بدون نیاز به کارهای میدانی اضافی نمونه‌برداری، مقادیر مناسبی از نمونه‌های آموزشی با کیفیت بالا را فراهم سازد [۷]. محدوده تالابی مورد مطالعه یکی از مهمترین و غنی‌ترین زیستگاه‌های مانگرو در ایران است. تالاب‌ها به‌دلیل خدمات حیاتی برای انسان و محیط زیست، یکی از مهمترین بوم‌سازگان‌ها هستند و نقشه‌برداری و پایش تالاب‌ها جهت حفاظت از آن‌ها ضروری است [۸].

در بیشتر مطالعات به تهیه نقشه‌های کاربری اراضی با الگوریتم‌های طبقه‌بندی مختلف پیکسل پایه و شی‌گرا و ارزیابی آن‌ها بر اساس معیارهای ارزیابی صحت و دقت نقشه‌های تولیدی توجه شده است و کمتر مطالعاتی به تهیه نمونه‌های آموزشی پویا برای تهییه نقشه‌های کاربری اراضی به‌روز، بر مبنای یک کاربست و چهاچوب کاربردی دائمی توجه شده است. در همین راستا، قربانیان و همکاران (۲۰۲۰)

1. Google Earth Engine
2. Google Earth Engine

با استفاده از روش نمونه‌های آموزشی متغیر پیکسل پایه و نیز شی‌گرا برای کل کشور ایران موفق به تهیه نقشه کاربری اراضی سال ۲۰۱۹ (سال مقصد) بر اساس نمونه‌های آموزشی سال ۲۰۱۷ با دقت صحت کلی ۹۱/۳۵ و ضریب کاپای ۹۱ درصد شدند. نامبردگان روش مذکور را برای تهیه نقشه‌های کاربری اراضی برای سایر مناطق نیز پیشنهاد دادند. ایشان در طبقه‌بندی تصاویر از روش جنگل تصادفی[1] (RF) استفاده کردند [۹]. همچنین فکری و همکاران (۲۰۲۱) با استفاده از شاخص‌های NDVI، NDWI و SD با استفاده از تصاویر سنتینل ۲ و سنتینل ۱ و الگوریتم طبقه‌بندی جنگل تصادفی و روش نوآورانه نمونه‌های آموزشی متغیر که با روش‌های فاصله زاویه طیفی و فاصله اقلیدسی نسبت به تصاویر مرجع سال ۲۰۱۸ برای سال‌های ۲۰۱۹، ۲۰۲۰ و ۲۰۲۱ تهیه کردند به پایش تغییرات کاربری اراضی تالاب شادگان در بستر GEE پرداختند [۸]. هوآنگ و همکاران (۲۰۱۵) نیز با انتخاب پیکسل‌هایی از هر کلاس و تهیه امضاهای طیفی از آنان، به احصاء روشی خودکار برای تهیه نمونه‌های آموزشی خودکار از پوشش زمین پیشنهاد دادند [۱۷]. این روش‌ها بدلیل اعمال فیلترهای مکانی و طیفی برای حذف پیکسل‌هایی که در امتداد مرزهای بین دو نوع پوشش زمین متفاوت قرار دارند، حذف مقادیر پراکنده از توزیع امضاهای طیفی و تغییر مداوم در کلاس‌های پوشش زمین قابل‌اعمال برای تحلیل‌های چندزمانه نیستند. صادقی و همکاران (۱٤۰۰) با استفاده از روشی نوآورانه از آزمون خی دو و خوشه‌بندی K_means با پالایش نمونه‌های آموزشی قدیمی به تهیه نقشه کاربری اراضی پرداختند. نامبردگان از طبقه‌بند ماشین‌بردار پشتیبان بهینه شده با الگوریتم ژنتیک استفاده کردند. در روش ایشان، انتخاب نمونه‌های آموزشی نزدیک به مراکز اصلی خوشه‌های داخلی هر کلاس با اعمال تنوع طیفی داخل کلاس‌ها لحاظ شدند [۲۷]. از طرفی، محمدی و خدابنده‌لو (۱۳۹۹) با اخذ تصاویر لندست ۸ در بستر GEE به طبقه‌بندی تصاویر اخذ شده با استفاده از روشهای ماشین‌بردار پشتیبان، جنگل تصادفی و حداقل فاصله از میانگین (روش‌های پیکسل پایه) و طبقه‌بندی شی‌گرا (نزدیکترین همسایه) پرداختند. نامبردگان روش ماشین‌بردار پشتیبان و شی‌گرا را دارای دقت بیشتری جهت طبقه‌بندی تصاویر ارزیابی کردند. نامبردگان در پژوهش خویش از شاخص‌هایی نظیر NDBI و NDVI و ... به‌منظور کاهش اختلاط کلاس‌ها و افزایش صحت طبقه‌بندی استفاده کردند [۲۵]. محمدنژاد آروق (۱۳۹۹) نیز با استفاده از شاخص‌های NDVI (میانگین پوشش گیاهی) و MNDWI (میانگین پوشش آب) بر مبنای تصاویر سنتینل ۲ در بستر GEE به شناسایی اراضی شهری پرداخت [۲٦].

در پژوهش حاضر هدف علاوه بر برآورد دقت طبقه‌بندی بر پایه روش‌های مختلف یادگیری ماشین، استفاده از شاخصه اهمیت نسبی باندها و مؤلفه‌های ورودی جهت تشخیص ویژگی‌های تصویر و عارضه بوده است. بنابراین از روش‌های یادگیری ماشینی استفاده شد که خروجی آن‌ها اهمیت نسبی داشته باشد. بنابراین روش‌هایی مانند ماشین بردار پشتیبان یا نزدیکترین همسایه از روش‌های بکار رفته در این پژوهش حذف شدند و سه روش جنگل تصادفی، رگرسیون تقویت شده[2] و درختان تصمیم چندگانه جمع‌شدنی[3] استفاده شدند. همچنین آزمودن روش نمونه‌های تعلیمی متغیر از سال مبدأ برای تهیه نمونه‌های تعلیمی سال‌های مقصد و برآورد دقت نقشه‌های سال‌های مذکور بر پایه روش فاصله زاویه طیفی بدون تهیه نمونه‌های تعلیمی برای تصاویر سال‌های مقصد از دیگر اهداف و نوآوری‌های پژوهش حاضر بوده است.

مواد و روش‌ها

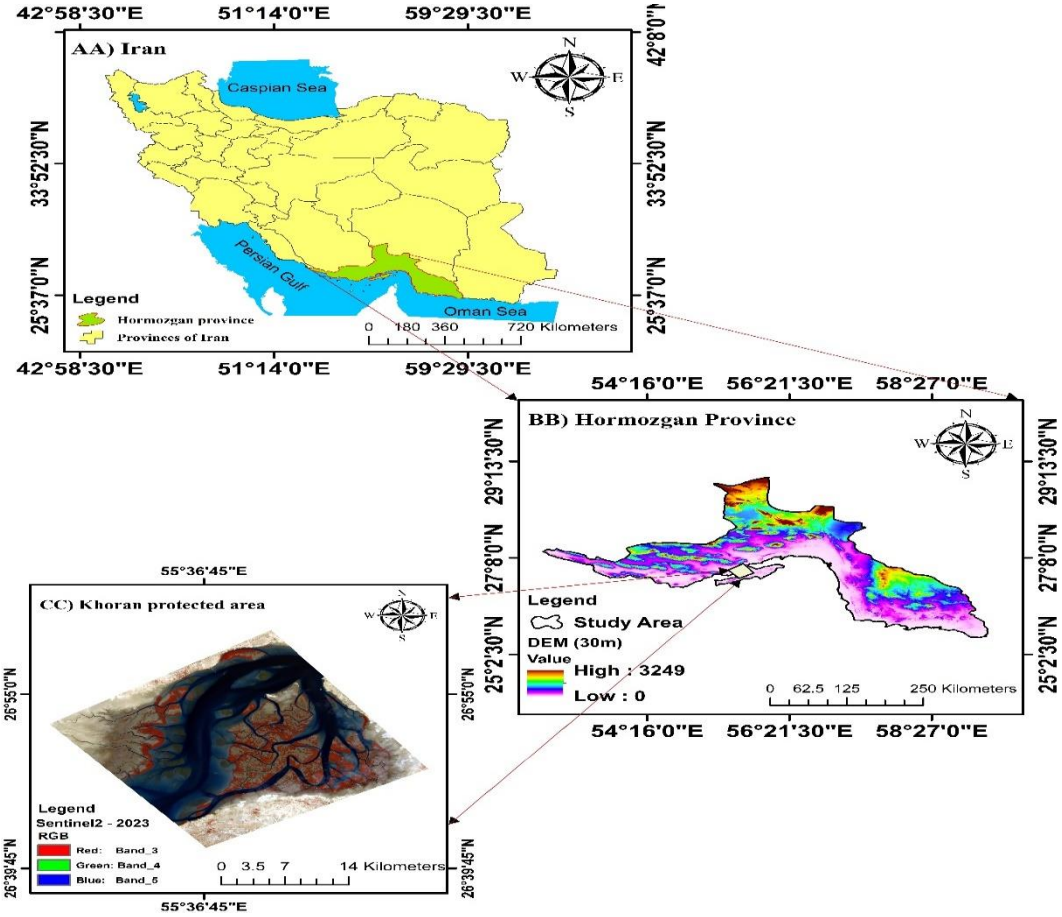
محدوده مورد مطالعه

منطقه حفاظت شده حرا و تالاب بین‌المللی خورخوران (ذخیره‌گاه زیست‌کره حرا)، یکی از جنگل‌های پنج‌گانه حرا در استان هرمزگان در شمال‌غربی جزیره قشم و جنوب شهر بندرخمیر است. این تالاب بزرگ‌ترین جنگل دریایی مانگرو در خاورمیانه است. این بخش از کشور جزء رویشگاه سودان – دکانی قرار دارد و جوامع حرا (*Avicennietum*) تنها اجتماعات مانگرو در این محدوده هستند. این جوامع خالص، تنک تا انبوه در نواحی کم‌شیب جزر و مدی در تماس با آب‌های با شوری بالا در شمالی‌ترین گستره مداری خود در آسیای غربی جای گرفته‌اند [٦]. متوسط بارندگی منطقه مورد مطالعه ۲۰۰ میلی‌متر است که قسمت اعظم آن طی ماه‌های بهمن و اسفند دریافت می‌شود. اقلیم منطقه گرم بیابانی و حداقل دما به طور متوسط ۱۸ درجه سانتی‌گراد (دی ماه) و حداکثر آن بطور متوسط ۳٤ درجه سانتی‌گراد (تیرماه) است. شکل ۱ موقعیت قرارگیری این محدوده را از سواحل جنوبی شهر خمیر تا روستای لافت و طبل جزیره قشم در موقعیت جغرافیایی '٤۳ و °۲٦ تا '۵۹ و °۲٦ عرض شمالی و '۲۷ و °۵۵ تا '٤۸ و °۵۵ طول شرقی را نشان می‌دهد.

داده‌های مورد استفاده در پژوهش حاضر تصاویر سنتینل ۲ بازتاب سطح یک با ۱۳ باند طیفی در سامانه گوگل ارت انجین فراخوانی شدند. باندهای ۲، ۳، ٤، ۷، ۸، ۱۱ و ۱۲ در این پژوهش استفاده شدند (باندهای آبی، قرمز، سبز، مادون قرمز نزدیک،SWIR1 و SWIR2 و RE[4] استفاده شدند). تمامی باندها با استفاده از عملیات هم‌پیکسل‌سایز (ریسمپل[5]) به ابعاد ۱۰ متری تبدیل شدند. نقشه مرجع برای تهیه داده‌های آموزشی مرجع از نقشه کاربری اراضی سازمان فضایی اروپا (ESA/WorldCover/v100) با قدرت تفکیک مکانی ۱۰ متر استفاده شد. در این رابطه ٦ کلاس کاربری اراضی شامل جنگل‌های حرا (کد صفر)، پهنه‌های آبی (کد یک)، شوره‌زارها و اراضی بایر (کد دو)، مناطق مسکونی و شهری (کد سه)، مراتع و زمین‌های کشاورزی (کد چهار) و اراضی مرطوب تالابی (کد

1. Random Forest

2. Gradient boosted tree
3. Classification and regression tree (CART)
4. Red Edge
5. Resampling



شکل ۱- موقعیت منطقه مورد مطالعه در کشور و استان هرمزگان

Fig 1. Location of the studied area in the Iran and Hormozgan province

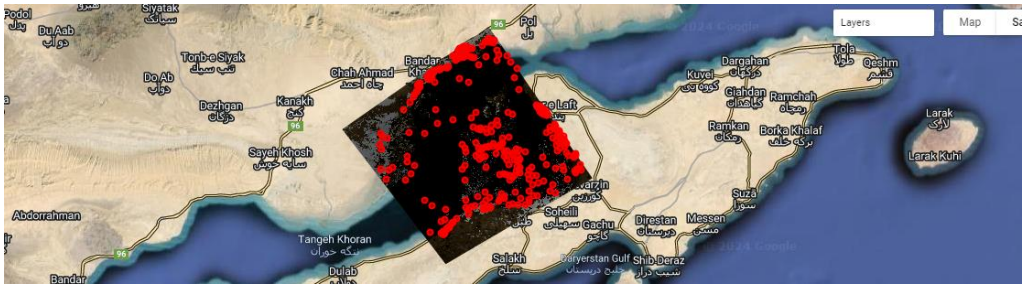
پنج) تعریف شدند. شایان ذکر است تصاویر ماهواره سنتینل ۲ در مرحله رشد فنولوژیک جنگل‌های حرا محدوده مورد مطالعه اخذ شدند. بعد از فراخوانی این تصاویر، تصاویر اخذ شده در زمان‌های مختلف با یکدیگر موزاییک شدند و سپس میانه این تصاویر به عنوان ورودی به فرآیند مدل‌سازی استفاده شد. در این پژوهش سعی شد علاوه بر استفاده از باندهای مختلف در طبقه‌بندی‌های پیکسل پایه از شاخص‌های طیفی بعنوان داده‌های کمکی در فرآیند طبقه‌بندی استفاده شود. قابل ذکر است تمامی تصاویر اخذ شده با اعمال تصحیح رادیومتریک و نیز حذف تصاویر بالای ۱۰ درصد پوشش ابری گرفته شدند. در مرحله بعد بر اساس نمونه‌های تعلیمی تصاویر موزاییک شده سال‌های مبدأ و مقصد، فاصله زاویه طیفی[1] (SAD) محاسبه و با اعمال آستانه‌گذاری، پیکسل‌های با احتمال بالای تغییر حذف و نمونه‌های آموزشی با کیفیت بالا، مناسب و بدون تغییر را برای سال هدف تهیه شد. برای محاسبه تفاوت بین طیف‌های مرجع و هدف و شناسایی نمونه‌های آموزشی ثابت، از شاخص فاصله زاویه طیفی (SAD) استفاده شد [۸، ۹] که در رابطه (۱) مشاهده می‌شود.

$$SAD = \cos^{-1}\theta = \frac{\cos(\theta)\ \sum_{i=1}^{N} A_{i(t1)} B_{i(t2)}}{\sqrt{\sum_{i=1}^{N}(A_{i(t=1)})^2 \sum_{i=1}^{N}(B_{i(t1)})^2}} \qquad (۱)$$

که در رابطه بالا، θ زاویه طیفی، A طیف مرجع در زمان t1، B طیف در زمان t2، i شاخص باند و N تعداد باندها است. محدوده SAD بین ۰ و ۱ است. اگر هر دو طیف مرجع و هدف یکسان باشند، مقدار SAD برابر یک است. لازم به ذکر است که اگرچه SAD بزرگی و جهت تغییر را نشان می‌دهد، تصاویر تغییر در سایر شاخص‌های کاربردی (یعنی NDVI، NDWI و SD) با تفریق مقادیر آن‌ها در هر جفت سال مرجع-هدف محاسبه می‌شوند.

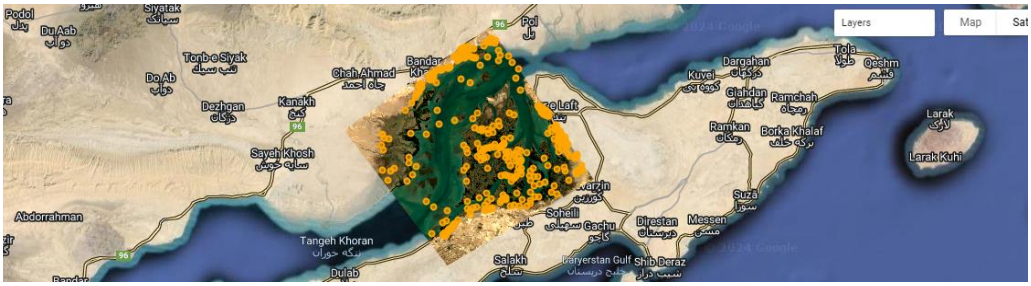
در ادامه کار، ۷۰ درصد نمونه‌ها برای آموزش و ۳۰ درصد داده برای تست و ارزیابی صحت و دقت نقشه‌های تولیدی استفاده شدند [۸]. برای جلوگیری از تعداد نمونه‌های ناکافی در تصاویر مقصد دو راهکار پیشنهاد شده است که در مرحله اول تعداد نمونه‌های آموزشی کافی از تصاویر مبدأ تهیه شود و در مرحله دوم در صورتیکه تعداد نمونه‌های مقصد کم باشند، به مجموعه داده‌های مذکور تعدادی نمونه آموزشی اضافه شود. شایان ذکر است کلیه موارد فوق در بستر GEE انجام گرفت. حدود آستانه‌گذاری با روش آزمون و خطا جهت انتقال نمونه‌های با کیفیت (بدون تغییر) انجام می‌گیرد. قربانیان و

1. Spectral Angle Distance (SAD)



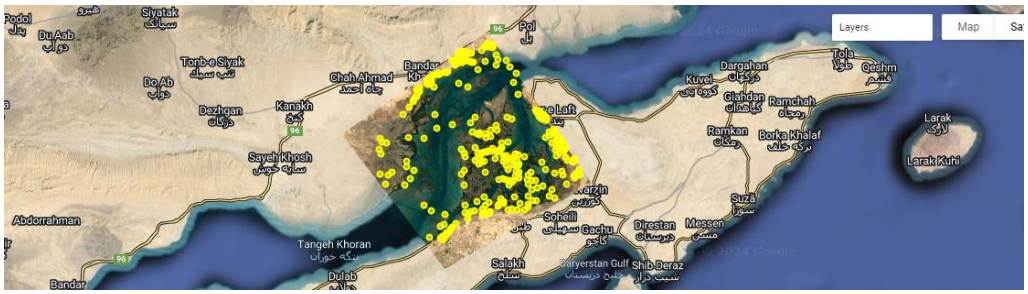
شکل۲- نمونه‌های آموزشی تصویر مرجع (۲۰۱۹) در سامانه گوگل ارت انجین

Fig 2. Training samples of reference image (2019) in the Google Earth Engine



شکل۳- نمونه‌های آموزشی تصویر مقصد (۲۰۲۲) در سامانه گوگل ارت انجین

Fig 3. Training samples of the destination image (2022) in the Google Earth Engine



شکل٤- نمونه‌های آموزشی تصویر مقصد (۲۰۲٤) در سامانه گوگل ارت انجین

Fig 4. Training samples of the destination image (2024) in the Google Earth Engine

همکاران (۲۰۲۰) حد این آستانه را ۰/۱ پیشنهاد دادند [۹]. فکری و همکاران (۲۰۲۱) حد این آستانه را ۰/٦ پیشنهاد دادند [۸]. در شکل‌های ۲ تا ٤ نمونه‌های آموزشی تصویر مرجع، مقصد ۲۰۲۲ و ۲۰۲٤ ارائه شده است.

همچنین استفاده از شاخص‌های مختلف طیفی مانند NDVI و NDWI و ... در شناسایی اجزای مختلف یک محدوده تالابی و تغییرات آن بسیار مفید هستند [۱۵]. در این مرحله ۱۱ شاخص مختلف به عنوان داده‌های کمکی[1] جهت طبقه‌بندی تصاویر سنتینل ۲ در ترکیبات باندی قرار گرفتند. شاخص‌های طیفی استفاده شده به شرح ذیل بودند؛

شاخص تفاوت نرمال شده گیاهی (NDVI[2])

این معیار پرکاربردی برای تعیین کمیت سلامت و تراکم پوشش گیاهی است. NDVI مقداری بین ۱- و ۱ دارد و در نقاطی که پوشش گیاهی وجود نداشته باشد NDVI صفر و منفی است. شاخص گیاهی تفاوت نرمال شده (NDVI) با استفاده از معادله ۲ محاسبه می‌شود [۱۳].

$$NDVI = \frac{NIR - R}{NIR + R} \qquad (۲)$$

شاخص پوشش گیاهی پیشرفته (EVI[3])

این شاخص به تغییرات ساختاری تاج پوشش، از جمله شاخص سطح برگ (LAI)، نوع تاج پوشش و فیزیولوژی گیاه حساسیت نشان می‌دهد. EVI به دلیل توانایی حذف نویزهای پس‌زمینه و جو کاربرد زیادی دارد. شاخص پوشش گیاهی پیشرفته (EVI) با استفاده از معادله ۳ محاسبه می‌شود [۳۱].

1. Auxiliary data
2. Normalized Difference Vegetation Index
3. Enhanced Vegetation Index

$$EVI = \frac{2.5(NIR - R)}{NIR + C_1R - C_2R + L} \quad (۳)$$

شاخص پوشش گیاهی حرا (MVI[1])

با توجه به حضور وسیع جنگل‌های حرا در محدوده مورد مطالعه از شاخص پوشش گیاهی حرا (MVI) برای ترسیم وسعت حرا از تصاویر سنجش از راه دور استفاده شد. معادله شاخص پوشش گیاهی حرا (MVI) به شرح رابطه ٤ است [۳].

$$MVI = \frac{NIR - Green}{SWIR1 - Green} \quad (٤)$$

شاخص پوشش گیاهی حرا پیشرفته (EMVI[2])

همچنین برای تدقیق بهتر و مجدد پوشش گیاهی مانگرو در محدوده مورد مطالعه از شاخص EMVI بر اساس باندهای سبز و مادون قرمز موج کوتاه استفاده شد و تفاوت در میزان سبزی و رطوبت تاج بین حرا و سایر پوشش‌های گیاهی تشخیص داده شد که رابطه ۵ نحوه محاسبه این شاخص را نشان می‌دهد [۳٤].

$$EMVI = \frac{Green - SWIR2}{SWIR1 - Green} \quad (۵)$$

شاخص گیاهی تعدیل کننده اثر خاک (SAVI[3])

شاخص گیاهی تنظیم شده خاک (SAVI) برای کاهش اثر تفاضل پوشش گیاهی نرمال شده (NDVI) بر تأثیر روشنایی خاک در مناطقی که پوشش گیاهی کم است استفاده می‌شود. برآورد این شاخص به شرح رابطه ٦ است [۱۳].

$$SAVI = (1 + L)\frac{NIR - R}{SWIR1 + R + L} \quad (٦)$$

شاخص گیاهی مقاوم در برابر جو (ARVI[4])

این شاخص مبتنی بر پوشش گیاهی است که اثرات پراکندگی جوی ناشی از ذرات معلق در هوا مانند باران، مه، گرد و غبار، دود یا آلودگی هوا را به حداقل می‌رساند. محاسبه شاخص (ARVI) به شرح رابطه ۷ است [۳۱].

$$ARVI = \frac{NIR + 2R + B}{NIR + 2R - B} \quad (۷)$$

شاخص نرمال شده رطوبت (NDWI[5])

این شاخص برای ترسیم ویژگی‌های آب مطرح شده است این شاخص با به حداکثر رساندن ویژگی‌های آب محاسبه می‌شود. محاسبه این شاخص به شرح رابطه ۸ است [۲۰].

$$NDWI = (Green - NIR)/(Green + NIR) \quad (۸)$$

شاخص نرمال شده ساخت و ساز (NDBI[6])

یکی از گسترده‌ترین شاخص‌ها جهت بررسی شرایط ساخت و ساز و مناطق مسکونی از مناطق غیرمسکونی است. این شاخص با استفاده از انعکاس باند مادون قرمز نزدیک و باند مادون قرمز میانی محاسبه می‌شود. این شاخص مقادیری بین ۱ تا ۱- دارد و مقادیر بزرگ‌تر این شاخص حکایت از ساخت و ساز و مقادیر کوچک‌تر آن نشان‌دهنده پوشش گیاهی یا خاک بایر و لخت است که در رابطه ۹ نحوه محاسبه آن بیان شده است [۲۱].

$$NDBI = \frac{SWIR1 - NIR}{SWIR1 + NIR} \quad (۹)$$

شاخص خاک لخت یا بدون پوشش (BSI[7])

این شاخص برای جداسازی کاربری‌هایی نظیر پوشش گیاهی و مناطق ساخت‌ساز شهری یا مسکونی از خاک بدون پوشش استفاده می‌شود. خاک معمولاً بازتابندگی بیشتری در باندهای قرمز و مادون قرمز کوتاه در مقایسه با مادون قرمز نزدیک و باند آبی دارد. این شاخص مقادیری بین ۱ تا ۱- دارد و مقادیر بالا بیانگر خاک لخت و خشک و مقادیر کمتر و منفی آن بیانگر پوشش‌هایی نظیر آب و گیاه است. رابطه ۱۰ نحوه محاسبه این شاخص را ارائه می‌دهد.

$$BSI = \frac{(Red + SWIR1) + (NIR + Blue)}{(Red + SWIR1) - (NIR + Blue)} \quad (۱۰)$$

شاخص منطقه خشک‌شده و سوخته پوشش گیاهی (NBR[8])

این شاخص برای مشخص کردن پوشش گیاهی سوخته و خشک شده استفاده می‌شود. این شاخص برای ایجاد تمایز بین پوشش سوخته و پوشش گیاهی سالم توسعه یافته است. در مناطق سوخته (گیاه و خاک سوخته) بازتاب مادون قرمز نزدیک کاهش می‌یابد. مقادیر این شاخص بین ۱ تا ۱- است و مقادیر بزرگتر نشان‌دهنده پوشش گیاهی سالم و مقادیر منفی و کوچکتر حکایت از گیاه ناسالم و سوخته دارد. نحوه محاسبه این شاخص در رابطه ۱۱ آورده شده است.

$$NBR = \frac{NIR - SWIR2}{NIR + SWIR2} \quad (۱۱)$$

الگوریتم‌های یادگیری ماشینی طبقه‌بندی تصاویر

جنگل تصادفی (RF[9])

الگوریتم جنگل تصادفی یکی از بهترین الگوریتم‌های یادگیری است که اولین بار توسط لئوبرایمن و آدل کاتلر توسعه داده شد. این الگوریتم نوع توسعه یافته‌ای از مدل طبقه‌بندی و رگرسیون درختی است [٤]. الگوریتم جنگل تصادفی مبتنی بر روش ترکیب اطلاعات است که در آن تعداد زیادی درخت تصمیم ایجاد شده و سپس نتایج تمام درختان برای پیش‌بینی ترکیب می‌شوند. در همین راستا، استفاده از چندین درخت تصمیم‌گیری حساسیت و بی‌ثباتی مدل را کاهش

1. Mangrove Vegetation Index
2. Enhanced Mangrove Vegetation Index
3. Soil-Adjusted Vegetation Index
4. Atmospherically Resistant Vegetation Index
5. Normalized Difference Water Index
6. Normalized Difference Built-up Index
7. Bare Soil Index
8. Normalized Burn Ratio
9. Random Forest

داده و منجر به قدرت و دقت مدل می‌شود. درختان تصمیم‌گیری با در نظر گرفتن بردار ورودی، آن را با هر درخت در جنگل طبقه‌بندی کرده و خروجی، برچسب‌های کلاسی هستند که از اکثریت آرا دریافت شده است. درختان در این مدل با استفاده از روش بوت استرپ ایجاد شده و در ساخت هر درخت، تعدادی از متغیرهای ورودی به طور تصادفی به منظور جلوگیری و کاهش تطابق بین درخت‌های تصمیم شرکت خواهند کرد. از جمله مزیت‌های این روش برای ساخت درخت می‌توان به کاهش ناسازگاری در نتایج مدل و افزایش قابلیت پیش‌بینی و نیز عدم نیاز به کالیبراسیون اضافی، اشاره کرد [۵].

مدل درخت گرادیان تقویت شده (GBT[1])

این مدل جز یکی از بهترین الگوریتم‌های یادگیری ماشینی است و در بسیاری از مجموعه داده‌ها کاربرد داشته و دقت قابل‌قبولی را ارائه کرده است. الگوریتم افزایش بر این اساس است که هنگام پیش‌بینی موارد پیچیده، نتایج حاصله از ادغام تعمیم چندین درخت، برتر از نتیجه حاصله از یک درخت خاص خواهد بود. در این روش درخت‌های تصمیم می‌توانند با استفاده از ویژگی‌های داده‌های ارائه شده مقادیر هدف را پیش‌بینی کنند، به طوری که در هر مرحله هر درخت زیرمجموعه برای بار اول با استفاده از داده‌هایی که به اشتباه توسط درخت پیشین پیش‌بینی شده‌اند،آموزش داده می‌شوند. همین امر موجب تمرکز بیشتر مدل بر روی موارد پیچیده شده و بنابراین دقت روش افزایش پیدا می‌کند. درختان گرادیان تصمیم‌گیرنده، بر خلاف مدل‌های خطی مانند رگرسیون لجستیک، تغییرات غیرخطی و متقاطع را در ویژگی‌های ورودی بیشتر می‌توانند اعمال کنند. به ویژه به دلیل پیکربندی سلسله مراتبی، آن‌ها مقیاس‌پذیرتر، قوی‌تر در پیدا کردن نقاط پرت دور افتاده و دارای توانایی زیاد در شبیه‌سازی مرزهای تصمیم‌گیری غیرخطی به طور طبیعی هستند. این نوع مدل برای مدیریت ویژگی‌های آماری با ده‌ها دسته مناسب هستند [۲۹].

درخت رگرسیون و طبقه‌بندی (CART[2])

درخت تصمیم از روش‌های داده کاوی به شمار می‌آید که در دو دهه اخیر توسعه زیادی یافته است. از این فن می‌توان برای کشف و استخراج دانش از یک پایگاه داده و هم برای ایجاد مدل‌های پیش‌بینی استفاده نمود. درخت تصمیم همچنین یکی از ابزارهای قوی و متداول برای دسته‌بندی و پیش‌بینی است. درخت تصمیم قادر به تولید توصیفات قابل درک برای انسان، از روابط موجود در یک مجموعه داده بوده و می‌تواند برای وظایف دسته‌بندی و پیش‌بینی به کار می‌رود. این ساختار تصمیم‌گیری می‌تواند به شکل‌های ریاضی و محاسباتی که به توصیف دسته‌بندی و عام‌سازی یک مجموعه از داده‌ها کمک می‌کند نیز معرفی شود.

1. Gradient Boosting Trees
2. Classification and Regression Trees

در روش مبتنی بر درخت تصمیم، پیدایش هر درخت از دو مرحله تشکیل شده است الف: مرحله ایجاد و رشد درخت و ب: مرحله هرس درخت با هدف حداقل کردن خطای پیش‌بینی. تمامی الگوریتم‌های ایجاد درخت با نگرش بالا به پایین اجرا می‌شوند [۱۰، ۱٤، ۲٤].

ارزیابی دقت نقشه‌های کاربری اراضی

جهت ارزیابی صحت و دقت نقشه‌های کاربری اراضی تولید شده در سامانه GEE با استفاده از ماتریس ابهام[3] روش‌های صحت کلی[4]، ضریب کاپا[5]، دقت تولیدکننده[6]، دقت کاربر[7]، خطای اضافه[8] و خطای حذف شده [9]مطابق با روابط ۱۲ تا استفاده شده است؛

$$OA = {}^{1}/_{N \sum P_{ii}} \qquad (۱۲)$$

در این رابطه OA دقت کلی، N تعداد پیکسل‌های آزمایشی، P_{ii} جمع عناصر قطر اصلی ماتریس خطا هستند.

$$Kappa = \left({}^{P_o - P_c}/_{1 - P_c}\right) \times 100 \qquad (۱۳)$$

در این رابطه P_o درستی پیکسل‌های مشاهده شده (صحت مشاهده شده) و P_c درستی مورد انتظار(احتمال توافق) هستند. هر چه مقدار آن به سمت یک تمایل داشته باشد حکایت از دقت بیشتر طبقه‌بندی دارد.

$$PA = (\frac{t_a}{g_a}) \times 100 \qquad (۱٤)$$

در این رابطه PA درصد دقت کلاس a برای دقت تولیدکننده (احتمال اینکه یک پدیده خاص روی زمین روی نقشه نیز درست طبقه‌بندی شده است چقدر است)، T_a تعداد پیکسل‌های صحیح طبقه‌بندی شده به عنوان کلاس a و g_a تعداد پیکسل‌های کلاس a در واقعیت زمینی

$$UA = (\frac{t_a}{n_i}) \times 100 \qquad (۱۵)$$

در این رابطه UA درصد دقت کلاس a برای دقت کاربر (احتمال اینکه یک پیکسل که در یک طبقه معین قرار گرفته است در حقیقت بیانگر همان طبقه برروی زمین است یا خیر؟) و n_i تعداد پیکسل‌های کلاس a در نتیجه طبقه‌بندی است. خطای حذف شده و خطای اضافه شده مطابق با رابطه ۱٦ محاسبه می‌شود؛

$$C_a = 1 - UA \qquad (۱٦)$$
$$O_e = 1 - PA$$

در این روابط C_a خطای لحاظ شده و O_e خطای حذف شده هستند. دقت کاربر مکمل خطای اضافه است و دقت تولیدکننده

3. Confusion matrix
4. Overall Accuracy
5. Kappa coefficient
6. Producer Accuracy
7. User Accuracy
8. Commission
9. Omission

مکمل خطای حذف شده است. حدود قابل قبول شاخص‌های مذکور بالاتر از ۰/۷ (۷۰٪) می‌باشند که هرچه به سمت ۱ میل نماید، دقت طبقه‌بندی بیشتر است [۱۹].

نتایج و بحث

تحقیق حاضر با استفاده از تصاویر سری زمینی ماهواره‌ای سنتینل ۲ در منطقه خورخوران (حدفاصل جزیره قشم و بندرخمیر در استان هرمزگان) برای سال ۲۰۱۹ به عنوان سال مبدأ و تصاویر سال‌های ۲۰۲۲ و ۲۰۲٤ به عنوان سال‌های مقصد جهت تهیه نقشه‌های کاربری اراضی با استفاده از روش نمونه‌های آموزشی متغیر در بستر GEE هدف‌گذاری شدند. خلاصه نتایج ارزیابی دقت و صحت نقشه‌های کاربری اراضی تولید شده در جدول ۱ آورده شده است.

همانگونه که در جدول ۱ مشاهده می‌شود از میان الگوریتم‌های طبقه‌بندی، بیشترین میزان دقت برای صحت کلی تصاویر طبقه‌بندی شده سال ۲۰۲٤، مربوط به طبقه‌بندی جنگل تصادفی با دقت ۰/۹۱۰٤، بیشترین میزان ضریب کاپا برای همان سال مربوط به طبقه‌بندی جنگل تصادفی با مقدار ۰/۸۷٤۲، بیشترین دقت تولیدکننده برای سال ۲۰۲٤، مربوط به طبقه‌بندی جنگل تصادفی با دقت ۰/۷۷۸۱ و بیشترین دقت کاربر نیز برای همان سال مربوط به طبقه‌بندی جنگل تصادفی با مقدار ۰/۷۵۹٦ بوده است. نتایج نشان داد که برای نمونه‌های آموزشی متغیر با روش زاویه طیفی، بهترین روش طبقه‌بندی برای سال مقصد ۲۰۲٤ روش جنگل تصادفی بود. همچنین برای سال مقصد ۲۰۲۲ از میان الگوریتم‌های طبقه‌بندی، بیشترین میزان دقت برای صحت کلی تصاویر طبقه‌بندی شده سال ۲۰۲۲، مربوط به طبقه‌بندی جنگل تصادفی با دقت ۰/۸۹۵۵، بیشترین میزان ضریب کاپا برای همان سال مربوط به طبقه‌بندی جنگل تصادفی با مقدار ۰/۸۵۷۰، بیشترین دقت تولیدکننده برای سال ۲۰۲٤، مربوط به طبقه‌بندی جنگل تصادفی با دقت ۰/۸۳۵۱ و بیشترین دقت کاربر نیز برای همان سال مربوط به طبقه‌بندی جنگل تصادفی با مقدار ۰/۸۲۰۹ بوده است. همچنین نتایج نشان داد برای نمونه‌های آموزشی متغیر با روش زاویه طیفی، بهترین روش طبقه‌بندی برای سال مقصد ۲۰۲۲ نیز روش جنگل تصادفی بود. کمترین میزان خطای لحاظ شده برای سال مقصد ۲۰۲۲ مربوط به روش جنگل تصادفی با مقدار ۰/۱۷۹۱ و نیز کمترین میزان خطای حذف شده برای همان سال با روش جنگل تصادفی ۰/۱٦٤۹ در بین روش‌های طبقه‌بندی بوده است. همچنین برای سال مقصد ۲۰۲٤، کمترین میزان خطای لحاظ شده با مقدار ۰/۲٤۰٤ مربوط به روش جنگل تصادفی و نیز برای همان سال کمترین میزان خطای حذف شده مربوط به روش جنگل تصادفی با مقدار ۰/۲۲۱۹ بوده است. جهت بررسی میزان تأثیر باندها و نیز شاخص‌های طیفی در شناسایی ویژگی‌های مختلف جهت طبقه‌بندی تصادیر با الگوریتم‌های طبقه‌بندی از روش اهمیت نسبی استفاده شد. در اشکال ۵ و٦ اهمیت نسبی برای موارد مذکور در سال‌های مختلف آورده شده است.

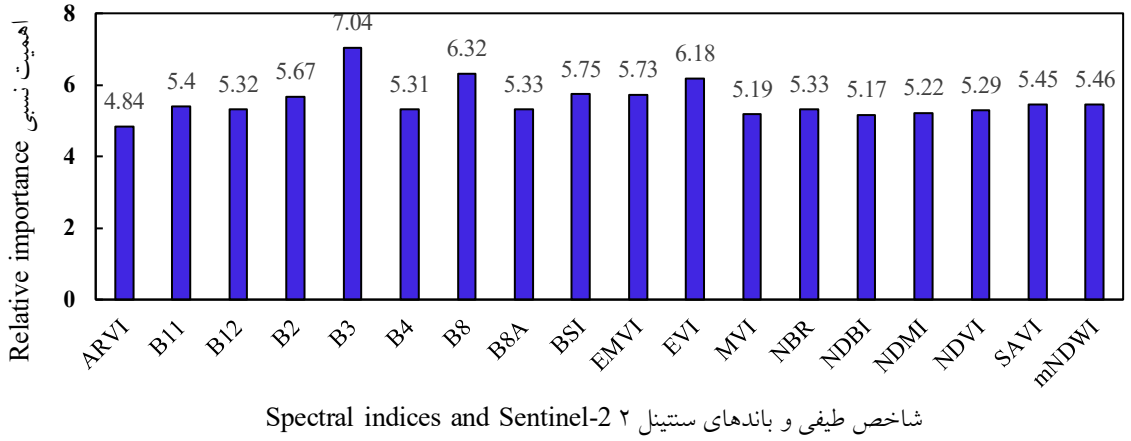
همانگونه که در شکل ۵ مشاهده می‌شود با استفاده از روش جنگل تصادفی برای سال ۲۰۱۹، مؤثرترین ویژگی‌ها برای طبقه‌بندی تصاویر سنتینل۲، شامل باندهای ۳، ۸، شاخص EVI بوده‌اند. لذا ترکیب بصری کاذب با ترکیب یادشده بیشترین تفکیک پدیده‌ها را بدنبال خواهد داشت.

همانگونه که در شکل ٦ مشاهده می‌شود با استفاده از روش جنگل تصادفی برای سال ۲۰۲۲، مؤثرترین ویژگی‌ها برای طبقه‌بندی تصاویر سنتینل۲، شامل باندهای ۱۱، A۸، شاخص EMVI، ۸ و NDVI و سپس EVI بوده‌اند. لذا ترکیب بصری کاذب با ترکیب یادشده خصوصاً با ترکیب باندهای ۱۱، A۸ و شاخص EMVI بیشترین تفکیک پدیده‌ها را بدنبال خواهد داشت. کمترین نقش در تفکیک پدیده‌ها در محدوده مورد مطالعه در سال ۲۰۲۲ مربوط به شاخص NDBI یا شاخص ساخت و ساز انسانی بوده است.

جدول۱- ارزیابی دقت و صحت نقشه‌های کاربری اراضی تولید شده در بستر GEE

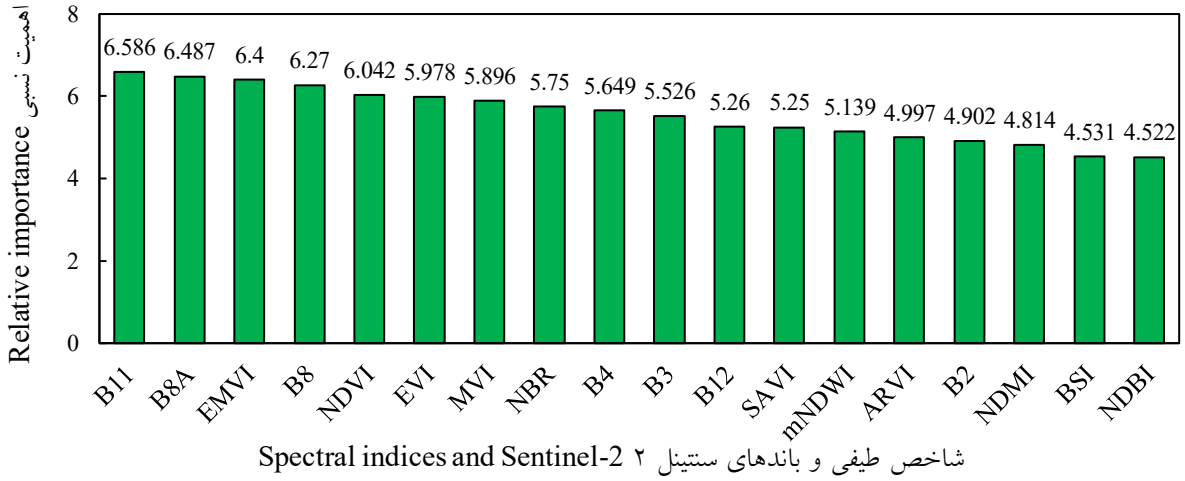
Table 1. Evaluation of the accuracy of land use maps produced in the GEE platform

ارزیاب Evaluator	سال و طبقه‌بندی RF Year and RF classification			سال و طبقه‌بندی GBT Year and GBT classification			سال و طبقه‌بندی CART Year and CART classification		
	2019	2022	2024	2019	2022	2024	2019	2022	2024
صحت کلی Overall accuracy	0.9391	0.8955	0.9104	0.9594	0.8507	0.8507	0.9391	0.7611	0.7611
ضریب کاپا Kappa coefficient	0.9212	0.8550	0.8742	0.9467	0.7939	0.7906	0.9214	0.6793	0.7175
دقت تولیدکننده Manufacture accuracy	0.8146	0.8351	0.7781	0.8530	0.7308	0.6825	0.8344	0.6862	0.7005
دقت کاربر User accuracy	0.8015	0.8209	0.7596	0.9583	0.7444	0.7317	0.8673	0.5652	0.6990
خطای لحاظ شده Error included	0.1985	0.1791	0.2404	0.417	0.2556	0.2683	0.1327	0.4348	0.3010
خطای حذف شده Error removed	0.1854	0.1649	0.2219	0.1470	0.2692	0.3175	0.1656	0.3138	0.2995



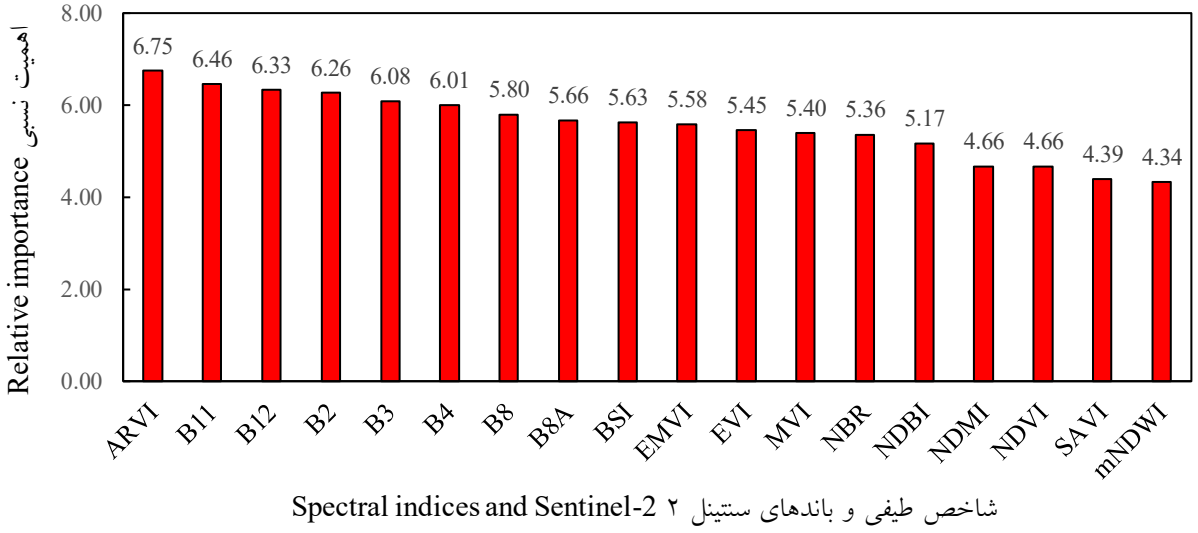
شکل ۵- اهمیت نسبی باندهای سنتینل۲ و شاخص‌های باندی در تشخیص ویژگی‌ها (جنگل تصادفی-۲۰۱۹)

Fig 5. The relative importance of Sentinel-2 bands and indices in feature selection (Random Forest - 2019)



شکل ٦- اهمیت نسبی باندهای سنتینل۲ و شاخص‌های باندی در تشخیص ویژگی‌ها (جنگل تصادفی-۲۰۲۲)

Fig 6. The relative importance of Sentinel-2 bands and indices in feature selection (Random Forest - 2022)



شکل ۷- اهمیت نسبی باندهای سنتینل ۲ و شاخص‌های باندی در تشخیص ویژگی‌ها (جنگل تصادفی-۲۰۲٤)

Fig 7. The relative importance of Sentinel-2 bands and indices in feature selection (Random Forest - 2024)

همانگونه که در شکل ۷ مشاهده می‌شود با استفاده از روش جنگل تصادفی برای سال ۲۰۲٤، مؤثرترین ویژگی‌ها برای طبقه‌بندی تصاویر سنتینل۲، شامل شاخص ARVI، باندهای ۱۱، ۱۲، ۲، ۳ و٤ بوده‌اند. لذا ترکیب بصری کاذب با ترکیب یادشده خصوصاً با ترکیب شاخص ARVI، باند۱۱ و ۱۲ بیشترین تفکیک پدیده‌ها را بدنبال خواهد داشت. کمترین نقش در تفکیک پدیده‌ها در محدوده مورد مطالعه در سال ۲۰۲۲ مربوط به شاخص mNDWI یا شاخص بهینه پهنه‌های آبی بوده است. همانطور که از شکل‌های ۵ تا ۷ مشخص است برای سال‌های مختلف باندها و شاخص‌های متفاوتی با توجه به شرایط نقش‌های پررنگ‌تر و کم‌رنگ‌تری در تفکیک پدیده‌ها و انتخاب ویژگی‌ها داشته‌اند. نتیجه اشتراکات و جمع‌بندی روش‌های مختلف طبقه‌بندی در سنوات مختلف و اهمیت‌نسبی باندها و شاخص‌ها برای آن سال‌ها نشان داد که شاخص‌های EMVI و mNDWI بدلیل غالب بودن پدیده‌های جنگل‌های مانگرو و نیز پهنه‌های آبی در محدوده مورد مطالعه غالبیت بیشتری از خود نشان داده‌اند و از بین باندهای سنجنده سنتینل ۲، باندهای ۳ ، ٤ و باند ۸ با اهمیت‌نسبی بیشتری خود را در تفکیک پدیده‌ها نشان دادند. بنابراین در ترکیبات باندی جهت تفکیک پدیده‌های مختلف استفاده از باندهای سبز، قرمز، مادون قرمز نزدیک و از میان شاخص‌ها، شاخص مانگرو و نیز شاخص بهینه پهنه‌های آبی برای محدوده مطالعاتی مناسب‌ترین تشخیص داده شدند و برای عرصه‌های مشابه در جنوب ایران و محدوده‌های مانگرو پیشنهاد می‌شوند.

بعد از مشخص شدن نقش باندها و شاخص‌های مختلف در تفکیک پدیده‌ها بر اساس اهمیت نسبی، بررسی تغییرات سطوح کاربری‌های مختلف با الگوریتم‌های طبقه‌بندی یادشده بررسی شد. در جداول ۳ میزان تغییرات سطوح در سنوات مختلف آورده شده است.

جدول ۲- باندها و شاخص‌های طیفی مؤثر در تفکیک ویژگی‌های تصاویر و طبقه‌بندی عارضه‌ها

Table 2. Effective spectral bands and indices for feature discrimination and object classification in imagery

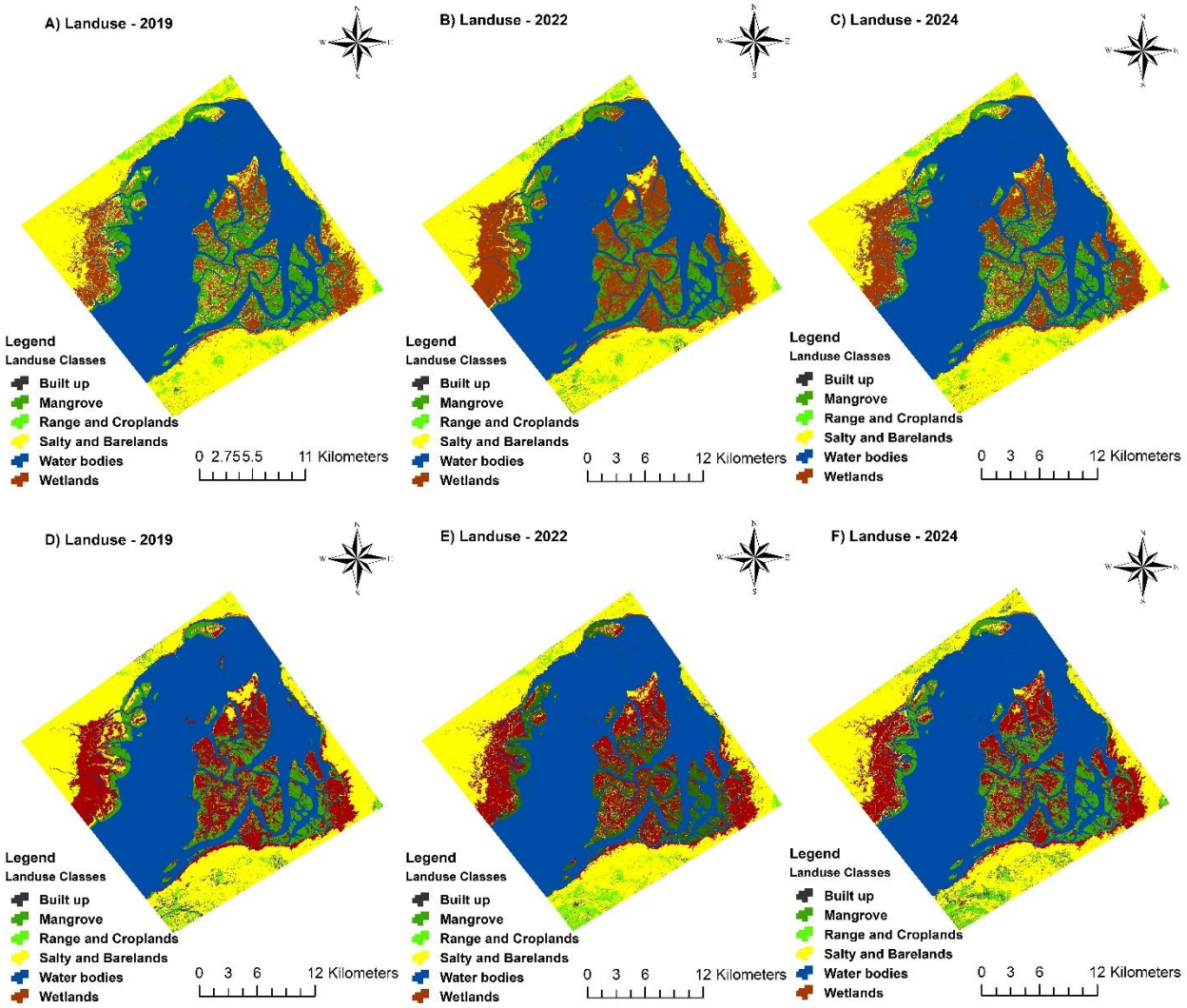
الگوریتم Algorithm	سال Year 2024	2022	2019
CART	B11-B8-B2-NDVI	mNDWI-NBR-EMVI-B8	B4-ARVI-EMVI-NDVI
GBT	mNDWI-B3-NBR-EMVI	B4-NBR-mNDWI-B3	mNDWI-B3-B4-B8
RF	B11-EVI-EMVI-B8A	B11-B8A-EMVI-B8	B3-B8-EVI-BSI
اشتراک‌ها Subscriptions	B11-EMVI	mNDWI-EMVI-B8	B3-B4
نتیجه‌گیری Conclusion	B3-B4-B8-B11-EMVI-mNDWI		

جدول ۳- تغییرات کاربری‌های اراضی بر اساس سال و مدل‌های طبقه‌بندی مختلف

Table 3. Land use changes based on year and different classification models

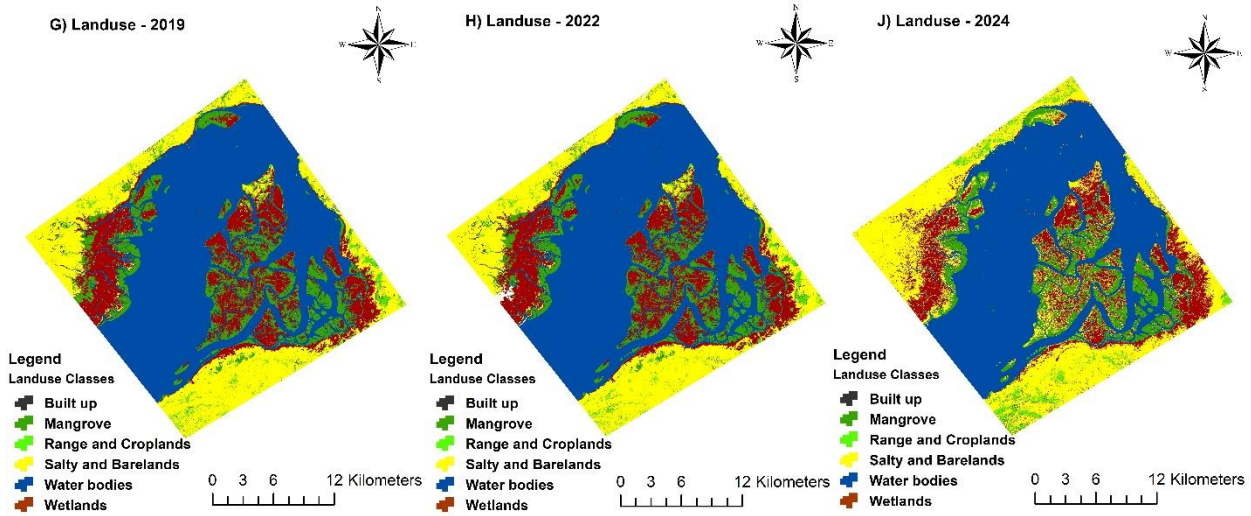
نوع کاربری Land use type	مساحت (هکتار) – سال (مدل جنگل تصادفی) Area (Ha)-Year (RF) 2024	2022	2019
Mangrove جنگل‌های حرا	6546.51	7037.65	7530.74
Water bodies پهنه‌های آبی	33741.20	32326.00	33133.30
Salty and Barelands شوره‌زارها و اراضی بایر	16431.50	15439.50	17729.90
Built up مناطق مسکونی و شهری	96.48	74.29	72.52
Range and Croplands زمین‌های کشاورزی	1211.82	1263.11	1456.42
Wetlands اراضی مرطوب تالابی	10526.40	12413.30	8651.03
نوع کاربری Land use type	مساحت (هکتار) – سال (مدل رگرسیون تقویت شده) Area (Ha)-Year (GBT) 2024	2022	2019
Mangrove جنگل‌های حرا	6571.54	6434.94	6710.93
Water bodies پهنه‌های آبی	33251.80	32547.30	31543.60
Salty and Barelands شوره‌زارها و اراضی بایر	15644.50	15943.10	16728
Built up مناطق مسکونی و شهری	648.16	331.05	811.291
Range and Croplands زمین‌های کشاورزی	1410.07	1667.75	756.49
Wetlands اراضی مرطوب تالابی	11027.70	11629.70	12003.50
نوع کاربری Land use type	مساحت (هکتار) – سال (مدل درختان چندگانه جمع‌شدنی) Area (Ha)-Year (CART) 2024	2022	2019
Mangrove جنگل‌های حرا	6369.74	7584.54	7798.56
Water bodies پهنه‌های آبی	32603.10	33803.00	33794.60
Salty and Barelands شوره‌زارها و اراضی بایر	18207.20	13777.20	13732.30
Built up مناطق مسکونی و شهری	214.36	233.67	232.84
Range and Croplands زمین‌های کشاورزی	1745.13	2081.96	2077.58
Wetlands اراضی مرطوب تالابی	9414.33	10919.90	10918.10

براساس روش‌های ارزیابی صحت و دقت نقشه‌های کاربری اراضی، روش جنگل تصادفی دارای بهترین دقت در میان سایر روش‌ها بود. مساحت جنگل‌های حرا در طول دوره بررسی از ۷۵۳۰/۷٤ هکتار به ٦٥٤٦/٥١ تنزل داشته است. چیزی در حدود ۹۸٤/۲۳ هکتار کاهش و به عبارتی ١٦٤ هکتار کاهش سطح جنگل‌های حرا به ازای هر سال. مساحت پهنه‌های مسکونی و عرصه‌های ساخت و ساز از ۷۲/۵۲ هکتار در سال ۲۰۱۹ به ۹٦/٤۸ در سال ۲۰۲٤ رسیده است که رشد شتابان این کاربری در دو سال آخر دیده می‌شود. همچنین کاهش اراضی کشاورزی و مرتعی در محدوده مورد مطالعه دیده شد، بطوریکه سطح این کاربری از ١٤٥٦/٤٢ هکتار در سال ۲۰۱۹ به ۱۲۱۱/۸۲ در سال ۲۰۲٤ تنزل یافته است. نواحی ماندابی و باتلاقی در محدوده مورد مطالعه افزایش سطح خصوصاً در سال ۲۰۲۲ نشان می‌دهد. وسعت این نواحی از ۸٦٥۱/۰۳ هکتار در سال ۲۰۱۹ به ۱۰۵۲۳/٤ هکتار در سال ۲۰۲٤ رسیده است. در شکل ۸ تغییرات کاربری اراضی با روش‌های طبقه‌بندی مختلف نشان داده شده است.



شکل ۸- A نقشه کاربری اراضی سال ۲۰۱۹ روش جنگل تصادفی، B نقشه کاربری اراضی سال ۲۰۲۲ روش جنگل تصادفی، C نقشه کاربری اراضی سال ۲۰۲٤ روش جنگل تصادفی، D نقشه کاربری اراضی سال ۲۰۱۹ روش رگرسیون تقویت شده، E نقشه کاربری اراضی سال ۲۰۲۲ روش رگرسیون تقویت شده، F نقشه کاربری اراضی سال ۲۰۲٤ روش رگرسیون تقویت شده

Fig 8. A land use map of 2019 using the RF method, B land use map of 2022 using the RF method, C land use map of 2024 using the RF method, D land use map of 2019 using the GBT method, E land use map of 2022 using the GBT method, F land use map of 2024 using the GBT method



ادامه شکل ۸- G نقشه کاربری اراضی سال ۲۰۱۹ روش درختان طبقه‌بندی و رگرسیون، H نقشه کاربری اراضی سال ۲۰۲۲ روش درختان طبقه‌بندی و رگرسیون ،J نقشه کاربری اراضی سال ۲۰۲٤ روش درختان طبقه‌بندی و رگرسیون

Fig 8. (continued). G land use map of 2019 using the CART method, H land use map of 2022 using the CART method, J land use map of 2024 using the CART method

الگوریتم‌های یادگیری ماشین مختلفی جهت طبقه‌بندی تصاویر ماهواره توسط محققین مختلف استفاده شده است. امانی و همکاران (۲۰۱۷) روش‌های نزدیکترین همسایه، بیشترین شباهت، ماشین بردار پشتیبان، درخت تصمیم و جنگل تصادفی را برای نقشه‌برداری مناطق مرطوب و تالابی استفاده کردند و در نهایت اعلام کردند دقت روش جنگل تصادفی نسبت به بقیه روش‌های یادگیری ماشین در طبقه‌بندی تصاویر بیشتر بوده است [۲]، در تحقیق حاضر نیز دقت روش جنگل تصادفی نسبت به دیگر روش‌های طبقه‌بندی تصاویر در سامانه گوگل ارت انجین مقادیر بیشتری از دقت و صحت را نشان داد. کارآمدی روش‌های نمونه‌های متغیر واقعیت زمینی یا نمونه‌های آموزشی متغیر برای تسری به سایر مناطق بدون نیاز به نمونه‌برداری واقعیت زمینی برای سال‌های مقصد نشان داده شده است [۹]، در تحقیق حاضر نیز سعی شد بر این مزیت تأکید گردد و نشان داده شود که با یکبار نمونه‌گیری برروی تصاویر مبدأ، طبقه‌بندی دقیق کاربری‌های اراضی برای سال‌های مقصد (هدف) با دقت مناسب محقق خواهد شد. فکری و همکاران (۲۰۲۱) اعلام کردند روش نمونه‌های آموزشی متغیر کارایی بالایی جهت تهیه نقشه‌های کاربری اراضی دارد [۸]؛ بطوریکه در تحقیق حاضر نیز کلیه شاخص‌های ارزیابی صحت و دقت نقشه‌های تولیدی برای سال‌های هدف (۲۰۲۲ و ۲۰۲٤) با روش جنگل تصادفی، مقادیر بالای ۰/۷ که حکایت از طبقه‌بندی مطلوب داشت را نشان دادند. این روش از مطالعات قبلی همچون جی و همکاران (۲۰۱۵) الهام گرفته شده است که در مطالعه یاد شده [۱۸]، در خصوص نمونه‌های آموزشی متغیر از اطلاعات طیفی مانند اندازه‌های شباهت به عنوان مبنایی برای محاسبه اندازه و جهت تغییرات بین طیف‌های مرجع و هدف (هوآنگ و همکاران، ۲۰۲۰؛ فان و همکاران،۲۰۲۱) استفاده شده بود [۱٦، ۲۷]. در مطالعه حاضر انتخاب محتمل‌ترین نمونه‌های آموزشی در دو سال هدف (۲۰۲۲ و ۲۰۲٤) با استفاده از روش نمونه‌های متغیر هدف بود با تکیه بر فاصله زاویه طیفی و نه تنها ویژگی‌های طیفی در نظر گرفته شد، بلکه شاخص‌های باندی و اهمیت نسبی آن‌ها در شناخت عوارض و تغییر از منظر مکانی با استفاده از اطلاعات بافت تصویر نیز مد نظر قرار گرفت که از این حیث مانند تحقیقات زانتوی و همکاران (۲۰۱۵) و مهدوی و همکاران (۲۰۱۷) و فکری و همکاران (۲۰۲۱) بوده است [۸، ۲۲، ۳۲]. علی‌رغم استفاده از داده‌های سری‌زمانی در تهیه نقشه‌های کاربری اراضی و استفاده از فضای ابری، پیچیدگی‌های محاسباتی بالا [۳۰] در مناطق وسیع همچون مقیاس کشور باعث شد که روش مذکور در مقیاس مطالعاتی کوچکتری مورد بررسی و اعتبارسنجی قرار گیرد. قربانیان و همکاران (۲۰۲۰) اعلام کردند، سامانه گوگل ارت انجین برای مناطق وسیع و با روش‌های شی‌گرا کارایی لازم را برای مناطق وسیع ندارد. بنابراین سعی شد اولاً در این تحقیق در یک مقیاس کوچک کارایی این سامانه بررسی شود و در ثانی بر خلاف نظر نامبردگان با روش‌های طیقه‌بندی یادگیری ماشینی به‌روز این سامانه می‌توان امکاناتی نظیر شاخص‌های طیفی، بافت، اهمیت‌نسبی و ... را در روند طبقه‌بندی و ارزیابی آن‌ها به کار گرفت. شایان ذکر است این سامانه محدودیت میلیون‌ها نمونه آموزشی را در ساختار خود دارد [۱]. فکری و همکاران (۲۰۲۱) با شاخص‌های صحت کلی، کاپا، دقت کاربر و تولیدکننده به ارزیابی نتیجه طبقه‌بندی با یک روش طبقه‌بندی (جنگل تصادفی) پرداختند اما در پژوهش حاضر علاوه بر روش‌های یاد شده از خطای لحاظ شده و خطای گماشته نیز استفاده شد و روش‌های طبقه‌بندی CART و GBT نیز بکارگرفته شد. همچنین در پژوهش یاد شده صرفاً از دو شاخص NDVI و

NDWI استفاده شده[۸] که در پژوهش حاضر تعداد قابل‌توجهی شاخص طیفی به عنوان داده کمکی به فرایند طبقه‌بندی تصاویر ماهواره‌ای در سامانه مذکور اضافه شد. قربانیان و همکاران (۲۰۲۰) صرفاً جهت ارزیابی نتایج طبقه‌بندی از دقت کلی و ضریب کاپا استفاده کردند و از شاخص‌های طیفی جهت طبقه‌بندی تصاویر سنتینل استفاده نکردند [۹]. پژوهش حاضر مانند تحقیق صادقی و همکاران (۱٤۰۰) جهت استخراج ویژگی از تصاویر و طبقه‌بندی آن‌ها شاخص‌های طیفی استفاده کرد [۲۷].

نتیجه‌گیری

توسعه ابزارهای مکانی از جمله سامانه برخط گوگل ارت انجین جهت مدیریت به‌موقع کاربری‌های اراضی از جمله پهنه‌های تالابی مانگرو ضروری است. بررسی مطالعات گذشته در طبقه‌بندی تصاویر ماهواره‌ای نشان داد که، یا از یک‌سری نمونه‌های تعلیمی (آموزشی) سال خاص و استفاده آن بدون تغییر برای سال‌های آتی بهره‌برداری می‌شود و یا برای هر سال یا هر تصویر نمونه‌های آموزشی جدیدی با مشقات فراوان تهیه می‌شود. از این رو، در پژوهش حاضر پیشنهاد شد که از یک کاربست ساده و مؤثر با توجه به نقاط واقعیت زمینی سال و تصاویر مبدأ و انتقال آن‌ها به عنوان نمونه‌های آموزشی در تصاویر سال‌های مقصد استفاده شود. در این پژوهش نمونه‌های آموزشی با کیفیت بالای ثابت از سال مرجع به سال هدف انتقال داده شد (مهاجرت نمونه‌های آموزشی از یک زمان خاص به سایر زمان‌ها) و دقت طبقه‌بندی بالایی با استفاده از الگوریتم طبقه‌بندی جنگل تصادفی به‌دست آمد. این روش در مطالعات چندزمانه کاربری‌های اراضی و کمبود یا ناکافی بودن نمونه‌های آموزشی یک راه حل بالقوه در سامانه گوگل ارت انجین است. پیشنهاد می‌شود در تحقیقات آینده از مجموع و برایند روش‌های ED[1]، SAD و K-means برای تولید نمونه‌های آموزشی متغیر استفاده و نتیجه طبقه‌بندی تصاویر با این روش‌ها مورد مقایسه و تحلیل قرار گیرد.

سپاسگزاری

نویسندگان از همراهی پژوهشکده هرمز دانشگاه هرمزگان تشکر می‌نمایند.

تضاد منافع نویسندگان

«نویسندگان این مقاله اعلام می‌دارند که هیچ‌گونه تضاد منافعی در خصوص نگارش و انتشار مطالب و نتایج این پژوهش ندارند.»

دسترسی به داده‌ها

«داده‌ها و نتایج استفاده شده در این پژوهش از طریق مکاتبه با نویسنده مسئول در اختیار قرار خواهد گرفت.»

مشارکت نویسندگان

محمد کاظمی: مفهوم‌سازی، انجام تحلیل‌های نرم‌افزاری، کنترل و نگارش نتایج و بدنه اصلی مقاله.

عاطفه جعفرپور: مفهوم‌سازی، ویرایش و بازبینی مقاله، بررسی چارچوب و نگارش مقاله.

منابع مورد استفاده

1. Amani, M. Brisco, B. Mahdavi, S. Ghorbanian, A. Moghimi, A. DeLancey, E. R…and LaRocque, A. 2020. Evaluation of the Landsat-based Canadian wetland inventory map using multiple sources: challenges of large-scale wetland classification using remote sensing. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 14: 32-52.

2. Amani, M. Mahdavi, S. and Berard, O. 2020. Supervised wetland classification using high spatial resolution optical, SAR, and LiDAR imagery. Journal of Applied Remote Sensing. 14(2): 024502-024502.

3. Baloloy, A. B., Blanco, A. C., Ana, R. R. C. S., & Nadaoka, K. (2020). Development and application of a new mangrove vegetation index (MVI) for rapid and accurate mangrove mapping. ISPRS Journal of Photogrammetry and Remote Sensing. 166, 95-117.

4. Breiman, L. 2001. Random forests. Machine learning. 45: 5-32.

5. Chen, W. Li, Y. Xue, W. Shahabi, H. Li, S. Hong, H. ... and Ahmad, B. B. 2020. Modeling flood susceptibility using data-driven approaches of naïve bayes tree, alternating decision tree, and random forest methods. Science of The Total Environment. 701: 134979.

6. Danehkar, A. and Jalali, S.Q. 2005, Avicennia marina forest structure using line plot method, Pajouhesh-va-sazandegi in natural resources. 67: 18-24. (In Persian)

7. Demir, B. Bovolo, F. and Bruzzone, L. 2012. Updating land-cover maps by classification of image time series: A novel change-detection-driven transfer learning approach. IEEE Transactions on Geoscience and Remote Sensing. 51(1): 300-312.

8. Fekri, E. Latifi, H. Amani, M. and Zobeidinezhad, A. 2021. A training sample migration method for wetland mapping and monitoring using sentinel data in google earth engine. Remote Sensing. 13(20): 4169.

9. Ghorbanian, A. Kakooei, M. Amani, M. Mahdavi, S. Mohammadzadeh, A. and Hasanlou, M. 2020. Improved land cover map of Iran using Sentinel imagery within Google Earth Engine and a novel automatic workflow for land cover classification using migrated training samples. ISPRS Journal of Photogrammetry and Remote Sensing. 167: 276-288.

10. Gomes Mantovani, R. Horváth, T. Rossi, A. L. Cerri, R. Barbon Junior, S. Vanschoren, J. and de Carvalho, A. C. 2024.

1. Euclidean Distance

Better trees: an empirical study on hyperparameter tuning of classification decision tree induction algorithms. Data Mining and Knowledge Discovery. 1-53.

11. Gorelick, N. Hancher, M. and Dixon, M. Ilyushchenko, S. Thau, D. and Moore, R. 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote sensing of Environment. 202:18-27.

12. Guo, M. Li, J. Sheng, C. Xu, J. and Wu, L. 2017. A review of wetland remote sensing. Sensors. 17(4): 777.

13. Halder, B. Bandyopadhyay, J. and Khatun, R. 2024. Google Earth Engine and Sentinel 1/2 data-based forest degradation monitoring of Sundarban Biosphere Reserve. Sustainable Horizons. 9: 100088.

14. Hasanah, U. and Nurhopipah, A. 2023. Prediksi Persentase Body Fat Menggunakan Algoritma CART dan M5'. JTIM: Jurnal Teknologi Informasi dan Multimedia. 4(4): 351-363.

15. Hu, T. and Smith, R. B. 2018. The impact of Hurricane Maria on the vegetation of Dominica and Puerto Rico using multispectral remote sensing. Remote Sensing. 10(6): 827.

16. Huang, H. Wang, J. Liu, C. Liang, L. Li, C. and Gong, P. 2020. The migration of training samples towards dynamic global land cover mapping. ISPRS Journal of Photogrammetry and Remote Sensing. 161: 27-36.

17. Huang, X.Weng, C. Lu, Q. Feng, T. and Zhang, L. 2015. Automatic labelling and selection of training samples for high-resolution remote sensing image classification over urban areas. Remote sensing. 7(12): 16024-16044.

18. Ji, L. Geng, X. Sun, K. Zhao, Y. and Gong, P. 2015. Target detection method for water mapping using Landsat 8 OLI/TIRS imagery. Water. 7(2): 794-817.

19. Kazemi, M. Mahdavi, Y. Nohagar, A. and Rezaei, P. 2011. Estimation of land cover and land use changes using remote sensing techniques and geographic information system (case study: Tang Bostank watershed, Shiraz), Journal of RS and GIS for natural resources. 2(1): 101-112. (In Persian)

20. Khosravian, M. Entezari, A. Rahmani, A. and Baaghide, M. 2017. Monitoring the Disturbance of Lake District Water Level Changes Using Remote Sensing Indices. Hydrogeomorphology. 4(13): 99-120. (In Persian)

21. Li, G. Meng, L. Zhang, M. and Wang, L. 2019. Mapping Urban Built-Up Areas Using Convolutional Neural Networks and MNDBI from Landsat 8 Imagery. Remote Sensing. 11(13): 1513.

22. Mahdavi, S. Salehi, B. Amani, M. Granger, J. E. Brisco, B. Huang, W. and Hanson, A. 2017. Object-based classification of wetlands in Newfoundland and Labrador using multi-temporal PolSAR data. Canadian Journal of Remote Sensing. 43(5): 432-450.

23. Mahdavi, S. Salehi, B. Granger, J. Amani, M. Brisco, B. and Huang, W. 2018. Remote sensing for wetland classification: A comprehensive review. GIScience & remote sensing. 55(5): 623-658.

24. Mishra, A. Singh, A. Tripathi, A. Pandey, S. K. and Srivastava, C. 2024. A study of deep learning, Twitter mining and machine learning based system for predicting customer churn. Artificial Intelligence, Blockchain, Computing and Security. 2: 461-469.

25. Mohammadi, A. and Khodabandehlou, B. 2020. Classification and Assessment of Land Use Changes in Zanjan City Using Object-Oriented Analysis and Google Earth Engine System. Geography and environmental planning. 31 (2): 25-42. (In Persian)

26. Mohammadnejad, V. 2020. Urban lands Extraction from Sentinel 1 and 2 satellite imagery based on Google Earth Engine (GEE). Geographical urban planning research. 8(3): 613-630. (In Persian)

27. Phan, D. C. Trung, T. H. Truong, V. T. Sasagawa, T. Vu, T. P. T. Bui, D. T. ... and Nasahara, K. N. 2021. First comprehensive quantification of annual land use/cover from 1990 to 2020 across mainland Vietnam. Scientific reports. 11(1): 9979.

28. Sadeghi, V. Ebadi, H. and Moghimi, A. 2021. Automatic Land Use/Land Cover Change Detection from Multitemporal Remote Sensed Images and Old Maps by Refining of Training Data Based on Chi-Square Test and K-Means Clustering. Journal of Geomatics Science and Technology. 10(4): 143-161.

29. Shadkani, S. Abbaspour, A. Samadianfard, S. Hashemi, S. Mosavi, A. and Band, S. S. 2021. Comparative study of multilayer perceptron-stochastic gradient descent and gradient boosted trees for predicting daily suspended sediment load: The case study of the Mississippi River, US. International Journal of Sediment Research. 36(4): 512-523.

30. Shaharum, N. S. N. Shafri, H. Z. M. Ghani, W. A. W. A. K. Samsatli, S. Prince, H. M. Yusuf, B. and Hamud, A. M. 2019. Mapping the spatial distribution and changes of oil palm land cover using an open source cloud-based mapping platform. International journal of remote sensing. 40(19): 7459-7476.

31. Somvanshi, S. S. and Kumari, M. 2020. Comparative analysis of different vegetation indices with respect to atmospheric particulate pollution using sentinel data. Applied Computing and Geosciences. 7: 100032.

32. Szantoi, Z. Escobedo, F. J. Abd-Elrahman, A. Pearlstine, L.

Dewitt, B. and Smith, S. 2015. Classifying spatially heterogeneous wetland communities using machine learning algorithms and spectral and textural features. Environmental monitoring and assessment. 187: 1-15.

33. Teimouri, S. Razavizadeh, S. Khosroshahi, M. Jalili, A. and Dargahian. F. 2024. Estimating the water requirement of Hamoons with the basis area derived from the Landsat product on the Google Earth Engine platform. Iranian journal of watershed management science. 18 (64) : 103-113. (In Persian)

34. Yang, G. Huang, K. Sun, W. Meng, X. Mao, D. and Ge, Y. 2022. Enhanced mangrove vegetation index based on hyperspectral images for mapping mangrove. ISPRS Journal of Photogrammetry and Remote Sensing. 189: 236-254.

نشریه علمی
علوم و مهندسی آبخیزداری ایران
Iran-Watershed Management
Science & Engineering

Vol. 18, No. 67, Winter 2025 سال هجدهم- شماره ۶۷- زمستان ۱۴۰۳

Multi Temporal Land use/Land Cover Using Spectral Indices, Sentinel-2 Imagery and Migrated Training Samples in Google Earth Engine

Mohamad Kazemi* [1]and Atefeh Jafarpoor[2]

Received: 24-07-2024 Accepted: 03-09-2024

Extended Abstract

Introduction

Remote sensing has become a valuable tool for acquiring integrated spatial data regarding land cover and land use across various temporal and spatial scales. One of the primary challenges in multi-temporal land cover and land use mapping is the availability and integrity of training data for supervised classification algorithms. Collecting training samples for each class of land cover and land use over different time periods can be time-consuming and, particularly in rapidly changing environments, fieldwork can be challenging. This issue is further exacerbated by the potential spectral and phenological changes in land cover features over time, which can diminish the transferability of training samples. The concept of "migration" or "transfer" of training samples from a reference year to target years has been explored as a means to overcome the limitations of training data. In this context, the use of Google Earth Engine (GEE) has facilitated multi-temporal land cover and land use mapping. GEE's ability to integrate diverse data sources, including Sentinel-2 imagery and a wide range of spectral indices, enables researchers to develop robust and scalable applications for land cover and land use classification. Monitoring changes in land use and cover over time is crucial for understanding and managing the environment. However, when there are limitations in training data for different time periods, this can pose significant challenges. This study presents an innovative approach to classify Sentinel-2 satellite images from different years using a set of reference training samples.

Materials and Methods

In this study, an innovative application was explored using migrated training samples from a reference year (Sentinel-2 imagery from 2019), along with bands from Sentinel-2 images and spectral indices, to classify land cover and land use in the dynamic and ecologically significant mangrove region (Khoran Protected Area). The use of machine learning algorithms within the Google Earth Engine (GEE) framework was examined to achieve high classification accuracy and monitor land cover changes over time. Satellite images from Sentinel-2 covering the study area for the target years 2022 and 2024, as well as the reference year 2019, were retrieved in GEE. Ground truth data, including the location and classification of various land cover types for the reference year, were collected using the European Space Agency's land use maps. Subsequently, high-quality ground truth data and their corresponding image samples from the reference year (2019) were transferred to the target year images using the Spectral Angle Distance (SAD) algorithm. Classification algorithms such as Random Forest (RF), Gradient Boosting Trees (GBT), and

1. Corresponding author and Associate Professor, Department of Hormoz Studies and Research Center, University of Hormozgan, Bandarabaas, Iran, mohamad.kazemi86@gmail.com
2.Assistant Professor, Faculty of Natural Resources and Environment, Shahid Bahonar University of Kerman, Kerman, 7616913439, Iran; Atefeh.jafarpoor@yahoo.com

Classification and Regression Trees (CART) were employed to classify the target year images using the variable training samples. The classified images were evaluated using various accuracy metrics, including overall accuracy, Kappa coefficient, producer accuracy, user accuracy, as well as commission and omission errors. Finally, the importance of different spectral bands from Sentinel-2 and spectral indices in the classification process was analyzed to identify the most suitable features for distinguishing various phenomena in the study area.

Results and Discussion

The results of this study indicated that among the classification algorithms, the highest accuracy for overall accuracy and Kappa coefficient in the classified images for 2024 and 2022 was achieved using the Random Forest classification, with accuracies of 0.9104 and 0.8742, and Kappa values of 0.8955 and 0.8570, respectively. Additionally, the findings revealed that the area of mangrove forests decreased from 7530.74 hectares to 6546.51 hectares during the study period, representing a reduction of approximately 984.23 hectares, or about 164 hectares per year.

The area of residential and construction zones increased from 72.52 hectares in 2019 to 96.48 hectares in 2024, indicating rapid growth in these land uses over the last two years. The synthesis and summary of various classification methods across different years highlighted the relative importance of bands and indices. Specifically, the EMVI and mNDWI indices demonstrated greater dominance due to their effectiveness in capturing the phenomena of mangrove forests and aquatic areas within the study region. Consequently, for the band combinations aimed at distinguishing various phenomena, the use of green, red, and near-infrared bands, along with the mangrove index and optimized water body index, were identified as the most suitable for the study area. These findings are recommended for similar areas in southern Iran and other mangrove regions. This research employed a simple and effective application by transferring ground truth points from the reference year to their corresponding images as training samples for the target year images, utilizing the Google Earth Engine platform. This approach holds the potential for extension to other regions.

Conclusion

In summary, this study demonstrates the potential of using migrated training samples and advanced machine learning algorithms such as Random Forest, Gradient Boosted Trees, and Classification and Regression Trees, along with spectral indices as auxiliary data, for the accurate classification of multi-temporal satellite imagery. The development of spatial tools, including the online Google Earth Engine (GEE) platform, is essential for the up-to-date management of land uses, particularly in wetland and mangrove areas. In this research, high-quality training samples were successfully migrated from the reference year to the target year, resulting in high classification accuracy using the Random Forest classification algorithm compared to other methods such as boosted regression or regression and classification trees. This method offers a viable solution in multi-temporal land use studies, especially in cases of insufficient or inadequate training samples within the GEE system. For future studies, it is recommended to employ a combination of Euclidean Distance (ED), Spectral Angle Distance (SAD), and K-means clustering for generating variable training samples, and to compare and analyze the classification results obtained through these methods. This approach presents a promising solution for producing up-to-date land use/land cover maps, even in challenging environments with limited training data. The findings of this study can guide future research aimed at monitoring land cover, managing, and

effectively protecting valuable natural resources, such as mangrove forests, in other regions.

Keywords: *Pixel-based classification algorithms, Relative importance, Spectral Angle Distance, Land use, Mangrove, Migrated training samples.*

Acknowledgement

The authors are grateful for the support of Hormoz Studies and Research Center of Hormozgan University.

Conflicts of interest

The authors of this article declared no conflict of interest regarding the authorship or publication of this article.

Data Availability Statement

The datasets are available upon a reasonable request to the corresponding author.

Authors' contribution

Mohamad Kazemi: Conceptualization, performing software analysis, controlling and writing, editing and revising the manuscript.

Atefeh Jafarpoor: Conceptualization, writing the manuscript, editing and revision.