



کاربرد سنجش از دور و GIS در علوم محیطی، شماره ۱۲، دوره ۴، فصل پاییز سال ۱۴۰۳،

صص ۱۱۶-۱۴۰

Application of remote sensing and GIS in environmental sciences, Vol.4., No.12, Autumn 2024,
pp.116-140



ارزیابی و تحلیل توسعه شهری و تأثیر آن بر کاهش نفوذپذیری سطوح شهر تبریز با استفاده از الگوریتم یادگیری ماشین و داده‌های سنجش از دور

محمد پردل، بهزاد قیاسی*

کارشناس ارشد، مهندسی عمران گرایش مدیریت ساخت، گروه مهندسی عمران، دانشکده مهندسی، دانشگاه بوعلی سینا،

همدان، ایران

استادیار دانشکده محیط زیست دانشگاه تهران، ایران

چکیده

طبقه‌بندی بهینه‌ی تصاویر ماهواره‌ای می‌تواند استخراج اطلاعات مکانی - زمانی برای طبقه‌بندی کاربری و پوشش زمین را تسهیل کند. در این مطالعه، با تحلیل توسعه شهری و تغییر پوشش گیاهی شهر تبریز به بررسی تغییرات سطوح نفوذناپذیر پرداخته شده است. با استفاده از تصاویر لندست، سه الگوریتم مورد استفاده قرار گرفت و بر اساس نتایج، الگوریتم طبقه‌بندی جنگل تصادفی در پلتفرم گوگل ارث انجین با استفاده از تصاویر لندست ۸ به‌عنوان ورودی برای دوره ۱۳۹۲ تا ۱۴۰۲ به کار برده شد. علاوه بر این، تأثیر ترکیب‌های تصویری مختلف بر اساس چندین شاخص طیفی بر دقت طبقه‌بندی نهایی بررسی شد. نتایج نشان می‌دهد که استفاده از محصولات لندست (TOA) و پلتفرم گوگل ارث انجین می‌تواند نتایج سریع‌تر و دقیق‌تری برای طبقه‌بندی ایجاد کند. همچنین مقادیر هایپرپارامتر برای ۳ سال ۱۳۹۲، ۱۳۹۷ و ۱۴۰۲ که بالاترین دقت را در مجموعه داده آزمایشی به همراه داشت، به ترتیب ۲۸۰، ۱۸۰ و ۸۰ درخت انتخاب شدند. یافته‌ها نشان دهنده افزایش پوشش گیاهی در کنار گسترش تدریجی شهری است. با این حال، این رشد در درجه اول به ایجاد مناطق تفریحی بزرگ و درختکاری در حومه شهر نسبت داده شد. با این حال، پوشش گیاهی در مناطق شهری مرکزی یا راکد مانده یا در مناطق خاصی کاهش یافته است. با توجه به پیامدهای تغییر اقلیم و افزایش بارندگی‌های شدید، کمبود پوشش گیاهی در مناطق مرکزی ممکن است منجر به تشکیل سطوح غیرقابل نفوذ شود و خطر سیل‌های ناگهانی را تشدید کند که اهمیت نیاز به نتایج این تحقیق و اقدامات پیشگیرانه را نشان می‌دهد.

کلمات کلیدی: کاربری اراضی، گوگل ارث انجین، هایپرپارامتر تیونینگ، جنگل تصادفی، لندست

۱- مقدمه

تغییرات در کاربری زمین، مانند شهرنشینی، جنگل زدایی و تخریب زمین های زراعی، می تواند به طور قابل توجهی چرخه هیدرولوژیکی را با تغییر تعادل بین بارش، نفوذ، تبخیر و تعرق و رواناب تحت تأثیر قرار دهد (Sajikumar & Remya, 2015). با افزایش جمعیت و گسترش مناطق شهری، الگوهای کاربری اراضی تغییر می کنند و این تغییر می تواند بر میزان نفوذپذیری سطوح، اثرگذار باشد. نفوذپذیری سطوح که به توانایی زمین در جذب و انتقال آب به لایه های زیرزمینی اشاره دارد، یکی از مهم ترین عوامل تأثیرگذار بر فرآیندهای هیدرولوژیکی و نحوه تشکیل رواناب در مناطق شهری است (Ghiasi et al., 2022; Tang et al., 2024). تغییر کاربری اراضی از مناطق طبیعی و نیمه طبیعی به کاربری های مسکونی، تجاری و صنعتی می تواند سبب افزایش سطوح غیرقابل نفوذ مانند ساختمان ها، جاده ها و پارکینگ ها شود. این امر، منجر به افزایش روان آب های سطحی و خطرات مرتبط با آن مانند سیلاب های شهری شود (L. Li et al., 2021; Vafa & Seker, 2023). همچنین، کاهش نفوذپذیری سطوح می تواند بر تغذیه آب های زیرزمینی و چرخه هیدرولوژیکی تأثیر منفی بگذارد. برای این منظور، تأثیر تغییر کاربری زمین بر سطح نفوذپذیری توسط بسیاری از محققان مورد مطالعه قرار گرفته است.

دینگ و شی^۱ (۲۰۱۳) در پژوهشی به بررسی تأثیر تغییر کاربری و پوشش زمین بر دمای سطح در شهر پکن پرداخته اند. یافته های تحقیقاتی آنها ارتباط قوی بین تغییر کاربری زمین و تغییرات دمای سطح را نشان داد. این مطالعه دیدگاه های ارزشمندی را در مورد تأثیر تغییر کاربری زمین بر شرایط محیطی ارائه می دهد و بر نیاز به ارزیابی جامع تر از تأثیرات تغییر کاربری زمین بر نفوذپذیری سطح تأکید می کند. در مطالعه ای دیگر، جمالی^۲ (۲۰۱۹) هشت مدل یادگیری ماشین را برای نقشه برداری کاربری زمین با استفاده از لندست OLI در منطقه شمال ایران ارزیابی و مقایسه کرده است. این مطالعه اثربخشی مدل های یادگیری ماشین را در نقشه برداری دقیق تغییرات کاربری و پوشش زمین برجسته کرد که نشان دهنده کاربرد بالقوه این مدل ها در ارزیابی تأثیر توسعه شهری بر کاهش نفوذپذیری سطح است. علاوه بر این، ژانگ^۳ و همکاران (۲۰۲۱) رابطه بین تغییر کاربری

زمین و محیط حرارتی شهری را با استفاده از الگوریتم‌های یادگیری ماشین در شهر ووهان چین بررسی کرده‌اند. یافته‌های آنها اهمیت الگوریتم‌های یادگیری ماشین را در شبیه‌سازی تأثیر تغییر کاربری زمین بر شرایط حرارتی شهری نشان داد. این مطالعه به درک اثرات توسعه شهری بر کاهش نفوذپذیری سطح و تشکیل جزیره گرمایی شهری کمک می‌کند. همچنین، نوراللهی و کابلی^۱ (۲۰۲۱) با استفاده از مدل‌های یادگیری ماشینی، از جمله GARP، RF، MaxEnt و NB، تحقیقاتی را بر روی نقشه‌برداری خطر سیل شهری انجام دادند. بیکاسانا و همکاران (۲۰۲۳) در مطالعه‌ای به تجزیه و تحلیل تغییرات طولانی‌مدت در کاربری و پوشش زمین در مانیپور هند، با استفاده از الگوریتم‌های یادگیری ماشین و داده‌های سنجش از دور از سال ۲۰۰۰ تا ۲۰۲۲ پرداخته‌اند. در این تحقیق، کاهش قابل توجهی در پوشش جنگل در مانیپور به دلیل فعالیت‌های انسانی در دو دهه گذشته، با دقت طبقه‌بندی از ۹۳٫۳۲٪ تا ۹۴٫۶۱٪ با استفاده از روش جنگل تصادفی، مشاهده شد. در یک تحقیق جامع که توسط المدرسی الحسینی و همکارانش در سال ۱۳۹۳ انجام شد، به بررسی تغییرات کاربری اراضی منطقه همدان با استفاده از روش پرسپترون چندلایه (MLP) پرداخته‌اند. نتایج این مطالعه نشان داد که حدود ۸۰۰ هکتار زمین کشاورزی و پوشش گیاهی در محدوده همدان در بازه زمانی ۲۰۰۲ تا ۲۰۰۹ کاهش قابل توجهی داشتند. این کاهش، نشان‌دهنده تحول قابل توجه در الگوهای کاربری اراضی منطقه است و تأثیرات زیست‌محیطی و پیامدهای چنین تغییراتی را روشن می‌کند.

افزایش سطوح غیرقابل نفوذ از قبیل آسفالت و بتن بر شدت میزان نفوذپذیری خاک تأثیر می‌گذارد، به طوری که امکان نفوذ آب به لایه‌های زیرزمینی به شدت کاهش می‌یابد. این امر علاوه بر افت سطح آب‌های زیرزمینی، خطر سیل خیزی در شهر را نیز افزایش می‌دهد. همچنین کاهش سطوح نفوذپذیر می‌تواند بر میکرو اقلیم شهری نیز تأثیر منفی بگذارد. به طوری که احتمال افزایش دما و کاهش رطوبت در محیط‌های شهری افزایش می‌یابد که بر آسایش حرارتی شهروندان تأثیر نامطلوبی دارد (جعفری و همکاران، ۱۳۹۹). بنابراین مدیریت این مشکلات مستلزم اتخاذ رویکردهای مختلفی از قبیل توسعه فضای سبز شهری، استفاده از مصالح نفوذپذیر در معابر و پارکینگ‌ها و همچنین نفوذ آب باران به زمین است که می‌تواند در کاهش آثار سوء توسعه شهری بر نفوذپذیری خاک مؤثر باشد.

این یافته‌ها در مجموع به درک تأثیر توسعه شهری و تغییر کاربری زمین بر کاهش نفوذپذیری سطح کمک می‌کنند. علاوه بر این، تحقیقات نشان می‌دهند که روش‌های سنتی نقشه‌برداری و نظارت بر کاربری اراضی، اغلب بر تفسیر دستی عکس‌های هوایی یا تصاویر ماهواره‌ای تکیه می‌کنند که می‌تواند زمان‌بر باشد و صرفه

اقتصادی نیز نداشته باشد. در سال‌های اخیر، تکنیک‌های پیشرفته سنجش از دور و یادگیری ماشین به‌عنوان ابزارهای مفید جهت نقشه‌برداری در مقیاس بزرگ از کاربری زمین شهری ظاهر شده‌اند (Aryal et al., 2023; Avand & Moradi, 2021; Hashemi Aslani et al., 2022; Tassi & Vizzari, 2020). در نتیجه، ادغام الگوریتم‌های یادگیری ماشین و تکنیک‌های سنجش از دور یک رویکرد امیدوارکننده برای ارزیابی و تحلیل تأثیر توسعه شهری و تغییر کاربری زمین بر کاهش نفوذپذیری سطح فراهم می‌کند. براین اساس تحقیق حاضر باهدف بررسی تغییرات کاربری اراضی و توسعه مناطق شهری و تأثیر آن بر نفوذپذیری سطوح با استفاده از الگوریتم یادگیری ماشین و داده‌های سنجش از دور انجام شده است که این روش می‌تواند بشکل مفهومی در روند مدیریت شهری و ساخت و ساز پروژه‌های عمرانی تغییراتی بنیادی ایجاد نماید.

در سال‌های اخیر، برای غلبه بر چالش تجزیه و تحلیل داده‌های بزرگ در سنجش از دور، گوگل پلت فرم گوگل ارث انجین را برای پردازش حجم عظیمی از داده‌ها برای مدت طولانی توسعه داد. پروژه GEE در سال ۲۰۱۰ راه اندازی شد و به محبوب‌ترین پلت فرم تجزیه و تحلیل رصدهای زمینی تبدیل شد. GEE مجموعه داده‌های بسیاری که شامل مجموعه داده‌های خام و پیش پردازش شده و مدل‌های ارتفاعی می‌باشد، در بر می‌گیرد. GEE همچنین وسعت مکانی مختلف منطقه‌ای، ملی و جهانی را پوشش می‌دهد. در میان مجموعه داده‌های موجود در GEE، آرشیو تصاویر لندست یکی از متداول‌ترین محصولات مورد استفاده برای کاربردهای مختلف زراعی و زیست محیطی است.

این مطالعه نسبت به تحقیقات پیشین برای تهیه نقشه‌های کاربری اراضی از چند جنبه متفاوت است. در این تحقیق برای افزایش وضوح باندهای RGB در داده‌های لندست، از پردازنده بالای جو^۱ (TOA) استفاده شد زیرا در این نوع تصویر، سطح بالاتری از اصلاحات رادیومتریک (تصحیحات مربوط به اثرات اتمسفری و زاویه تابش خورشید) اعمال می‌شود. در حالیکه اکثر محققان از بازتاب سطحی^۲ (SR) به جای داده‌های TOA استفاده می‌کنند. همچنین، در بیشتر تحقیقات برای مطالعه توسعه شهری از شاخص طیفی NDBI استفاده می‌گردد. در حالیکه در این مطالعه از شاخص طیفی UAI نیز در الگوریتم جنگل تصادفی برای تعیین ویژگی‌های اساسی در فرآیند نقشه‌برداری کاربری اراضی استفاده گردید. نتایج نشان داد که این شاخص دقت نقشه‌برداری برای مطالعه توسعه شهری را افزایش می‌دهد. نوآوری بعدی این تحقیق، مربوط به مجموعه داده‌های

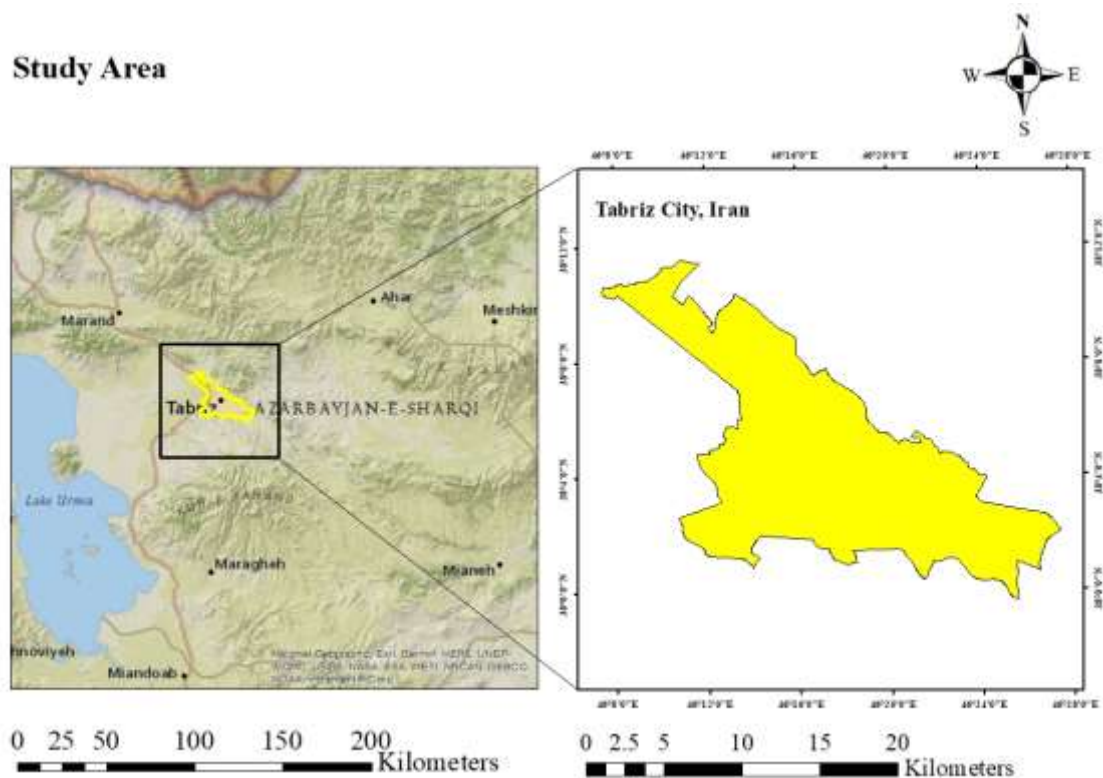
Google Earth Engine ^۱
Red Green Blue ^۲
Top Of Atmosphere ^۳
Surface reflection ^۴

آموزشی می باشد. برای افزایش دقت نقاط مورد اعتماد جهت آموزش ماشین، مشخصات مختصات نقاط در مرحله اول از طریق گوگل ارث استخراج گردید و بعد از مقایسه بصری با نقشه لایه‌های پوشش زمین جهانی کوپرنیک (CGLCL) به گوگل ارث انجین وارد شدند. سپس تمامی آن نقاط با توجه به پیکسل های موجود در لایه های NDVI، NDBI، MSAVI، UAI و True color بار دیگر مورد بازبینی قرار گرفته و تنها نقاطی به عنوان نقاط نهایی جهت آموزش و ارزیابی مدل انتخاب شدند که تمامی شرایط لازم را برای تمامی مراحل برآورده کرده اند. همچنین در این تحقیق از روش بهینه سازی فرارامتر تیونینگ جهت بهبود کیفیت مدل و دستیابی به نتایج بهتر در داده‌های آزمایشی استفاده گردید. استفاده از تعداد زیاد درخت، بدلیل پیچیده تر شدن محاسبات، مدت زمان انجام آن افزایش می‌یابد، اما نتایج حاصل از دقت بالایی برخوردار خواهد شد و حتی می تواند منجر به کاهش دقت گردد. این چارچوب دقت طبقه بندی را تصاویر لندست افزایش می دهد و امکان نظارت و تجزیه و تحلیل مداوم و طولانی مدت تغییرات کاربری را فراهم می کند.

۲- مواد و روش تحقیق

۲-۱- منطقه مورد مطالعه

شهر تبریز با مختصات جغرافیایی ۳۸ درجه ۰۵ دقیقه عرض شمالی و ۴۶ درجه ۱۷ دقیقه طول شرقی، در شمال غربی ایران قرار دارد و مرکز استان آذربایجان شرقی می‌باشد (شکل ۱). این شهر در شرقی ترین نقطه دشت دریاچه ارومیه قرار دارد که ارتفاع تقریبی آن از سطح دریا، ۱۳۰۰ تا ۲۱۰۰ متر متغیر می‌باشد و از همه جهات به جز شرق و شمال شرقی توسط کوهپایه های شیب دار احاطه شده است. میانگین دمای سالانه ۱۲٫۲ درجه سانتی‌گراد و میزان بارندگی سالانه ۳۱۱/۱ میلی متر می‌باشد که بیشینه بارش در زمستان و بهار ثبت گردیده است (بانان فردوسی و دین پژوه، ۱۳۹۸). به دلیل آب‌وهوای نیمه خشک شهر تبریز، این منطقه دارای تراکم کم پوشش گیاهی می‌باشد. پوشش گیاهی متراکم، محدود به دشت هایی است که آب های سطحی یا زیرزمینی در قالب مناطق کشاورزی در دسترس هستند (موسوی و یزدانی چهاربرج ۱۳۹۴). از نیمه دوم قرن بیستم، مهاجرت فزاینده و توسعه صنعتی، رشد جمعیت را تسریع بخشیده که برای یک دوره ۵۰ ساله (۱۳۸۴-۱۳۳۴) شهر تبریز دارای رتبه ی بالایی در کشور است (مرکز آمار ایران، ۱۴۰۱). جمعیت شهرستان تبریز در سال ۱۳۹۵ خورشیدی بالغ بر ۱،۷۷۳،۰۳۲ نفر بوده که از این تعداد ۱،۶۴۹،۲۵۸ نفر ساکن ۴ شهر تبریز، سردرود، خسروشاه و باسمنج و ۱۲۳،۷۷۴ نفر نیز ساکن روستاهای شهرستان بوده‌اند. امروزه تبریز، مرکز اصلی فرهنگ، صنعت، بازرگانی و حمل و نقل است. تبریز به‌عنوان بزرگ‌ترین شهر نیمه غربی کشور برای این پژوهش انتخاب شده است که رشد سریع جمعیت و گسترش شهری را نشان می‌دهد.



شکل ۱. محدوده منطقه مورد مطالعه

۲-۲- داده‌های مورد استفاده

در این مطالعه جهت بررسی توسعه شهری کلان‌شهر تبریز، تصاویر ماهواره‌ای مربوط به سنجنده TOA لندست ۸ برای سه دوره زمانی سال‌های ۱۳۹۲، ۱۳۹۷ و ۱۴۰۲ استفاده شد. از تصاویر مربوط به ماه جولای برای سال‌های ۲۰۱۳ و ۲۰۲۳ و از ماه آگوست برای سال ۲۰۱۸ استفاده گردید. زیرا در این ماه‌ها تراکم ابر کم می‌باشد و پوشش گیاهی به حداکثر رشد خود رسیده‌اند. مجموعه نمونه‌های آموزشی و اعتبارسنجی بر اساس تفسیر بصری تصاویر با وضوح بالا از گوگل ارث و نقشه CGLCL بدست آمدند. نقشه پوشش زمین با وضوح ۱۰۰ متر (CGLS-LC100) یک محصول جدید در مجموعه CGLS است و یک نقشه جهانی پوشش زمین را با وضوح مکانی ۱۰۰ متر ارائه می‌دهد (جدول ۱).

جدول ۱. داده های مورد استفاده در تحقیق

اسم تصویر	منبع تصویر	تاریخ قابل دسترس	رزولوشن
Landsat 8 Collection 2 Tier 1 calibrated top-of-atmosphere (TOA) reflectance	USGS/Google	۱۸-۰۳-۲۰۱۳ تا زمان حال	۱۵-۳۰ متر
CGLS-LC100 collection 3	Copernicus	۰۱-۰۱-۲۰۱۵ تا ۳۱-۱۲-۲۰۱۹	۱۰۰ متر

۳-۲- روش اجرای پژوهش

تغییرات کاربری زمین و پوشش سطح زمین در مقیاس‌های مختلف محلی، منطقه‌ای و جهانی بسیار پیچیده هستند و نیازمند درک عمیق‌تر فرایندها و نیروهای مؤثر بر این تغییرات هستند. ارزیابی تغییرات کاربری زمین طی زمان می‌تواند به درک بهتر فرایندهای تغییر کاربری و ارتباط آن با نیروهای محرک احتمالی کمک کند (Ren et al., 2019). این اطلاعات برای توسعه مدیریت زیست‌محیطی و اقدامات حفاظتی، پیش‌بینی کاربری آینده و مدل‌سازی اثر تغییرات کاربری زمین گذشته و آینده بر اجزای مختلف سیستم‌های طبیعی از جمله هیدرولوژی حوضه آبخیز یا تنوع زیستی ضروری است (Weber et al., 2001). برای اجرای این مطالعه، نیاز هست نقشه‌های کاربری اراضی تهیه گردد. در این تحقیق، از روش طبقه‌بندی نظارت شده بر روی پلتفرم گوگل ارث انجین استفاده شده است. گوگل ارث انجین یک پلتفرم محاسباتی قدرتمند است که مجموعه گسترده‌ای از الگوریتم‌های طبقه‌بندی پیشرفته را در خود جای داده است و توسط گوگل توسعه داده شده است (Yang et al., 2022). این پلتفرم به محققان و کاربران امکان دسترسی و تحلیل میلیاردها داده ماهواره‌ای و داده‌های زمینی را می‌دهد. گوگل ارث انجین دارای ویژگی‌های منحصربه‌فردی است که آن را به ابزار مناسبی برای تحقیقات در زمینه‌های مختلف محیط‌زیست تبدیل کرده است (Amani et al., 2020). همچنین دارای مخزن عظیمی از داده‌های ماهواره‌ای از جمله تصاویر ماهواره‌های لندست، سنتینل، مودیس است که به‌صورت رایگان در دسترس کاربران قرار دارد. این ابزار، یک رابط برنامه‌نویسی قدرتمند مبتنی بر جاوا اسکریپت و پایتون ارائه می‌دهد که امکان توسعه الگوریتم‌ها و اجرای تحلیل‌های پیچیده را فراهم می‌کند. به لطف محیط کاربری آنلاین و ابزارهای تعاملی موجود در گوگل ارث انجین، کاربران می‌توانند به‌سرعت وارد فرایند تحلیل داده شوند و نتایج را مشاهده کنند (Tamiminia et al., 2020).

طبقه‌بندی تصاویر ماهواره‌ای شامل تبدیل یک تصویر چند باندهای ماهواره‌ای به نقشه‌ای از کاربری‌های مختلف است (Palsson et al., 2011). در این فرایند، پیکسل‌های مشابه در تصویر به‌عنوان یک طبقه در نظر گرفته می‌شوند. تصاویر ماهواره‌ای ممکن است دارای بازتاب‌های متنوع در طیف امواج الکترومغناطیسی باشند، اما پیکسل‌هایی که به یک کاربری تعلق دارند در نقشه نهایی به‌عنوان یک طبقه نشان داده می‌شوند (Chapa et al., 2019). طبقه‌بندی به دو صورت نظارت شده و نظارت نشده انجام می‌شود. در نوع نظارت نشده، نرم‌افزار به طور خودکار منطقه را دسته‌بندی می‌کند، اما در نوع نظارت شده، کاربر نسبت به منطقه و باندهای تصویر آشنایی دارد (Egorov et al., 2015). در این مطالعه، تصاویر لندست ۸ با سه روش طبقه‌بندی نظارت شده K نزدیک‌ترین همسایه، ماشین بردار پشتیبان و جنگل تصادفی مورد ارزیابی قرار گرفت.

الگوریتم K نزدیک‌ترین همسایه، یک الگوریتم یادگیری ماشینی است که برای طبقه‌بندی و رگرسیون استفاده می‌شود. این روش بر اساس شباهت یا فاصله نمونه‌های آموزشی از نمونه پیش‌بینی شده کار می‌کند. این الگوریتم ساده و قابل فهم است و عملکرد خوبی در بسیاری از مسائل داده‌کاوی و یادگیری ماشینی دارد. ولی در مورد داده‌های پر ابعاد و پر نویز، ممکن است کارایی آن کاهش یابد. همچنین انتخاب مناسب مقدار K یک چالش است که باید با آزمایش‌های تجربی تعیین شود (Yu et al., 2018). الگوریتم بعدی، ماشین بردار پشتیبان (SVM) هست که از جمله الگوریتم‌های یادگیری ماشینی ناپارامتری است که برای انجام طبقه‌بندی کاربرد دارد. SVM با ایجاد یک هایپر صفحه (فضای تصمیم‌گیری) که بهترین جدایی بین دو کلاس را ایجاد می‌کند، عمل طبقه‌بندی را انجام می‌دهد (Karimi et al., 2019). ویژگی‌های اصلی SVM عبارت‌اند از: ۱. قابلیت طبقه‌بندی خطی و غیرخطی ۲. استخراج ویژگی‌های کرنل برای مسائل غیرخطی ۳. مقاومت در برابر پیش‌برازش با استفاده از حداکثر حاشیه‌ای ۴. بهینه‌سازی مسئله به‌صورت برنامه‌ریزی چندجمله‌ای ۵. قابلیت اعمال محدودیت‌های نامعادله برای ایجاد تصمیم‌گیری دقیق‌تر (Aryal et al., 2023). SVM در مسائل مختلفی از جمله تشخیص چهره، طبقه‌بندی متن، حشره‌شناسی، بیوانفورماتیک و غیره استفاده می‌شود (Salcedo-Sanz et al., 2014). این الگوریتم به‌طور گسترده‌ای در صنایع مختلف کاربرد دارد و در طول سال‌های اخیر توجه زیادی را به خود جلب کرده است.

الگوریتم بعدی استفاده شده در این تحقیق الگوریتم جنگل تصادفی است. این الگوریتم که یکی از الگوریتم‌های پیشرفته یادگیری ماشینی است، بر پایه دسته‌ای از درخت‌های تصمیم ساخته شده است. در این الگوریتم، تعداد

زیادی درخت تصمیم برای طبقه‌بندی رشد می‌کنند و پیکسل یا پدیده طبقه‌بندی نشده بر اساس ویژگی‌های خود در یکی از این کلاس‌ها قرار می‌گیرد. این کار با عبور دادن پیکسل از درختان تصمیم انجام می‌شود و هر درخت به کلاسی که به نظر خود مناسب است، رأی می‌دهد. در نهایت، پیکسل به کلاسی اختصاص می‌یابد که بیشترین رأی را از درختان جنگل دریافت کرده است (Belgiu & Drăguț, 2016). همچنین، درختان تصمیم‌گیری به صورت مستقل از مجموعه داده آموزشی رشد می‌کنند. در این فرایند، با نمونه‌برداری با جایگزینی دو سوم از داده‌ها برای آموزش هر درخت استفاده می‌شود و یک سوم باقی‌مانده به عنوان داده‌های آزمون مورد استفاده قرار می‌گیرند. این ویژگی باعث شده است که الگوریتم جنگل تصادفی در مورد مجموعه‌های داده‌های بزرگ نیز کارایی بالایی داشته باشد (Speiser et al., 2019).

جهت اعتبارسنجی الگوریتم‌های استفاده شده از روش دقت کلی و ضریب کاپا استفاده شد. دقت کلی از طریق نسبت مجموع عناصر قطر اصلی ماتریس خطا به تعداد کل پیکسل‌ها محاسبه می‌شود. این شاخص نشان‌دهنده میزان اطمینان نتایج به دست آمده از شاخص‌های کاربردی و گوگل ارث انجین در استخراج نقشه‌ها از تصاویر ماهواره‌ای هست. در این زمینه، میزان دقت کلی بیشتر یا مساوی ۸۵ درصد، قابل قبول است (Heydari & Mountrakis, 2018). ضریب کاپا (K) یک معیار برای ارزیابی توافق بین داده‌های مرجع (واقعیت) و نتایج طبقه‌بندی است. این ضریب تفاوت بین توافق واقعی و توافق تصادفی را نشان می‌دهد. هر چه ضریب کاپا بالاتر باشد، نتایج طبقه‌بندی به واقعیت نزدیک‌تر خواهد بود (Talukdar et al., 2020). برای انتخاب الگوریتم طبقه‌بندی تصاویر، این ضریب در کنار دقت کلی مدنظر قرار می‌گیرد تا بهترین عملکرد حاصل شود. ضریب کاپا از طریق رابطه ۱ محاسبه می‌شود (Foody, 2020):

$$k = \frac{N \sum_{i=1}^n m_{i,i} - \sum_{i=1}^n (G_i C_i)}{N^2 - \sum_{i=1}^n (G_i C_i)} \quad \text{فرمول (۱)}$$

که در آن، K ضریب کاپا، i شماره کلاس، N تعداد کل مقادیر طبقه‌بندی شده در مقایسه با مقادیر واقعی، $m_{i,i}$ تعداد مقادیر متعلق به کلاس واقعی i است که به عنوان کلاس i نیز طبقه‌بندی شده‌اند، C_i تعداد کل مقادیر پیش‌بینی شده متعلق به کلاس i و G_i تعداد کل مقادیر متعلق به کلاس i هستند.

داده‌های لندست دارای چندین باند مختلف بر اساس طول موج الکترومغناطیسی هستند (Gorji et al., 2020). به منظور تشخیص سریع و طبقه‌بندی دقیق اراضی ساخته شده با استفاده از تصاویر ماهواره‌ای، شاخص‌های طیفی مختلف ارائه شده است. برای این منظور در این مطالعه از شاخص پوشش گیاهی تفاضلی نرمال شده

(NDVI)، شاخص ساخت و ساز تفاضلی نرمال شده (NDBI)، شاخص منطقه شهرنشینی (UAI) و شاخص گیاهی اصلاح شده خاک (MSAVI) استفاده گردید که در ذیل به طور مختصر توضیح داده شده است. شاخص NDVI یکی از پرکاربردترین شاخص‌های پوشش گیاهی برای پایش سبزی‌نگی در سطح جهانی است. سایر شاخص‌های متداول پوشش گیاهی شامل شاخص پوشش گیاهی بهبودیافته (EVI)، شاخص پوشش گیاهی عمود (PVI)، و شاخص پوشش گیاهی نسبتی (RVI) هستند (Huang et al., 2021). به‌طور کلی، پوشش گیاهی سالم، جذب‌کننده خوبی در طیف مرئی است. کلروفیل موجود در گیاهان، جذب بالایی در باندهای آبی (۰,۴ تا ۰,۵ میکرون) و قرمز (۰,۶ تا ۰,۷ میکرون) دارد و در باند سبز (۰,۵ تا ۰,۶ میکرون) انعکاس زیادی دارد؛ بنابراین، چشم انسان پوشش گیاهی سالم و پرشتاب را به رنگ سبز می‌بیند. همچنین، گیاهان سالم انعکاس بالایی در باند مادون قرمز نزدیک (NIR) در محدوده ۰,۷ تا ۱,۳ میکرون دارند. این ویژگی به ساختار داخلی برگ‌های گیاهان مربوط است. این انعکاس بالا در NIR در مقابل جذب زیاد در باند قرمز، برای محاسبه شاخص NDVI مورد استفاده قرار می‌گیرد. فرمول محاسبه NDVI در رابطه ۲ آورده شده است (S. Li et al., 2021):

$$NDVI = \frac{(NIR-RED)}{(NIR+RED)} \quad \text{فرمول (۲)}$$

شاخص‌های بسیاری برای آنالیز مناطق ساخت و ساز شده در سنجش از دور وجود دارد. شاخص ساخت و ساز تفاضلی نرمال شده (NDBI) و شاخص منطقه شهرنشینی (UAI) از رایج‌ترین این شاخص‌ها برای آنالیز مناطق ساخت و ساز شده به حساب می‌آیند. مناطق ساخته شده و خاک برهنه، انعکاس مادون قرمز موج کوتاه (SWIR) بیشتر از مادون قرمز نزدیک (NIR) هست. مناطق آبی در طیف مادون قرمز انعکاس ندارند. در مورد سطوح سبز و پوشش گیاهی، انعکاس NIR بیشتر از طیف SWIR است (Zheng et al., 2021). شاخص NDBI و UAI را می‌توان با استفاده از رابطه‌های ۳ و ۴ محاسبه کرد (Bai et al., 2024; Guha et al., 2021):

$$NDBI = \frac{(SWIR-NIR)}{(SWIR+NIR)} \quad \text{فرمول (۳)}$$

$$UAI = BLUE + (2.5GREEN) - (1.5SWIR) - (0.25NIR) - (0.25RED) \quad \text{فرمول (۴)}$$

شاخص تمایز ساختمان‌ها (NDBI) دامنه‌ای از مقادیر -۱ تا +۱ را پوشش می‌دهد. در این دامنه، مقادیر منفی نشان‌دهنده حضور پهنه‌های آبی هستند، در حالی که مقادیر مثبت بیانگر وجود مناطق ساخت و سازی می‌باشند. همچنین، مقادیر مثبت UAI، معمولاً مناطق مسکونی یا شهری را نشان می‌دهد که NDBI بالاتر از NDVI

است و مقادیر منفی، مناطق دارای پوشش گیاهی یا غیر شهری را نشان می‌دهد که NDVI بالاتر از NDBI است.

شاخص گیاهی تعدیل شده با خاک (MSAVI) در جایی کار می‌کند که سایر شاخص‌های گیاهی این کار را نمی‌کنند. این شاخص را می‌توان در سنجش‌ازدور برای تشخیص رشد ناهموار بذر استفاده کرد. می‌توان آن را با داده‌های آب‌وهوا در نمودار مقایسه کرد که ارتباط بین آب‌وهوای شدید و سلامت محصول را آشکار می‌کند. از MSAVI برای تصحیح شاخص (NDVI) برای تأثیر روشنایی خاک در مناطقی که پوشش رویشی کم است استفاده می‌شود. فرمول این شاخص به صورت رابطه ۵ می‌باشد (Deng et al., 2023):

$$MSAVI = \frac{2NIR+1-\sqrt{(2NIR+1)^2-8(NIR-RED)}}{2} \quad \text{فرمول (۵)}$$

در این تحقیق، ۷۰ درصد نمونه‌ها بصورت تصادفی جهت آموزش و ۳۰ درصد آن‌ها جهت صحت سنجی استفاده شدند. با توسعه کد جاوا اسکریپت، تصاویر با چهار کاربری (پوشش گیاهی، مناطق شهری، آب و اراضی بایر)، توسط سه الگوریتم طبقه‌بندی نظارت شده k نزدیک‌ترین همسایه، ماشین بردار پشتیبان و جنگل تصادفی با پارامترهای پیش فرض، طبقه‌بندی شدند.

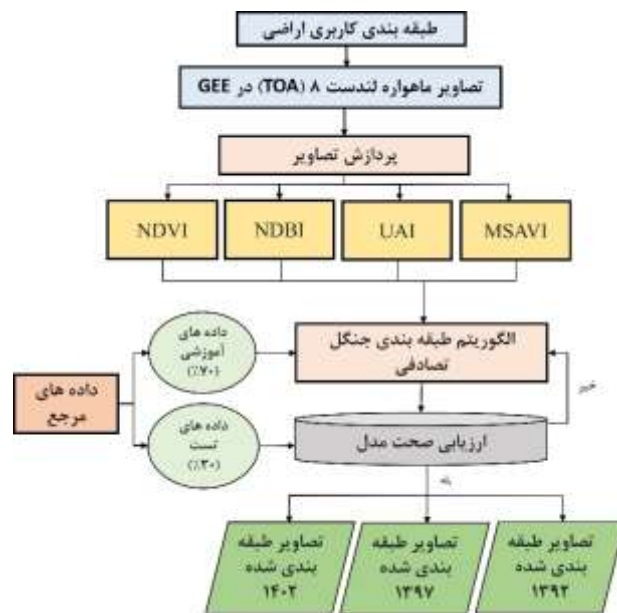
جدول ۲. مقایسه دقت طبقه‌بندی الگوریتم‌های طبقه‌بندی k نزدیک‌ترین همسایه، ماشین بردار پشتیبان و جنگل تصادفی

الگوریتم‌های طبقه‌بندی	جنگل تصادفی	ماشین بردار پشتیبان	k نزدیک‌ترین همسایه
دقت کلی آموزشی	۰/۹۹	۰/۸۳	۰/۷۶
دقت کلی صحت سنجی	۰/۹۶	۰/۸۱	۰/۷۳
ضریب کاپا	۰/۹۴	۰/۷۹	۰/۶۶

چنان که در جدول ۲ مشاهده می‌کنید، باتوجه به میزان کاپا و دقت کلی، الگوریتم جنگل تصادفی در مقایسه با سایر روش‌های طبقه‌بندی، از دقت بیشتری برخوردار بوده و بهترین نتیجه را ارائه داده است. بنابراین، در این مطالعه از الگوریتم جنگل تصادفی در پلتفرم گوگل ارث انجین، برای انجام طبقه‌بندی پوشش زمین استفاده شد. همچنین در این مطالعه، روش هایپرپارامتر تیونینگ برای بهینه‌سازی عملکرد مدل جنگل تصادفی برای

طبقه‌بندی کاربری و پوشش زمین مورد استفاده قرار گرفت. هدف، افزایش دقت مدل و دستیابی به نتایج بهتر در مجموعه داده‌های آزمایشی بود.

در انتها، دقت مدل آموزشی نهایی با استفاده از داده‌های اعتبارسنجی ارزیابی شد. برای ارزیابی دقت مدل طبقه‌بندی شده از دقت کلی^۱ (OA)، دقت تولیدکننده^۲ (PA)، دقت کاربر^۳ (UA)، ضریب کاپا^۴ (K) و امتیاز F1 استفاده گردید. شباهت‌های بین نمونه‌های طبقه‌بندی شده توسط مدل و داده‌های اعتبارسنجی به صورت کمی در ماتریس‌های خطا مقایسه شد. OA نسبت نمونه‌هایی است که با توجه به برچسب‌های ورودی آنها به درستی طبقه‌بندی شده‌اند. با این حال، ضریب کاپا نمایش مفیدتری از عملکرد است. زیرا OA تمایل دارد عملکرد واقعی را بیش از حد تخمین بزند. PA نسبت نمونه‌هایی است که به درستی در یک گروه دسته‌بندی شده‌اند به کل نمونه‌های مرجع در آن گروه، در حالی که UA نسبت نمونه‌هایی است که به درستی در یک گروه خاص دسته‌بندی شده‌اند به کل نمونه‌هایی که در آن گروه دسته‌بندی شده‌اند. شکل ۲ روند مراحل کار برای تولید نقشه کاربری اراضی را نشان می‌دهد.

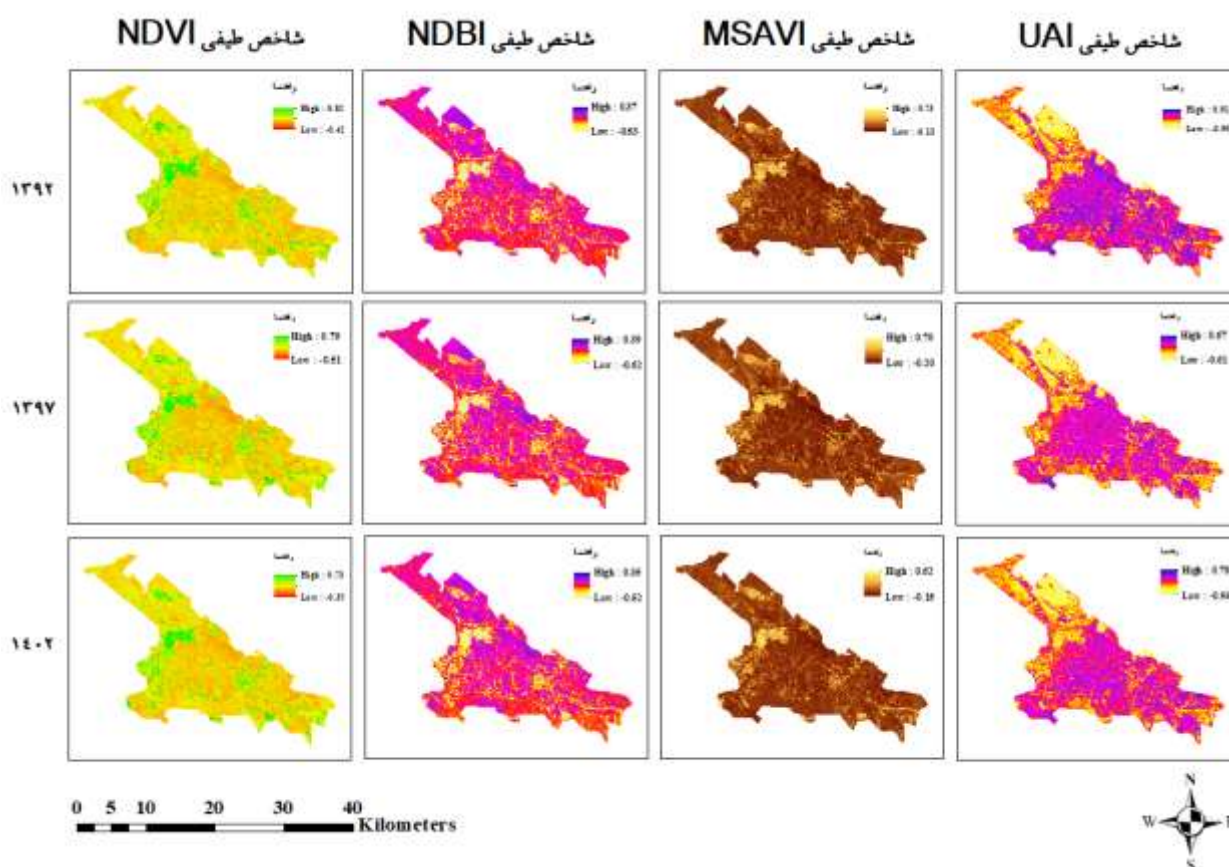


Overall Accuracy^۱
 Producer Accuracy^۲
 User Accuracy^۳
 kappa coefficient^۴

شکل ۲. فلوجارت روش انجام تحقیق

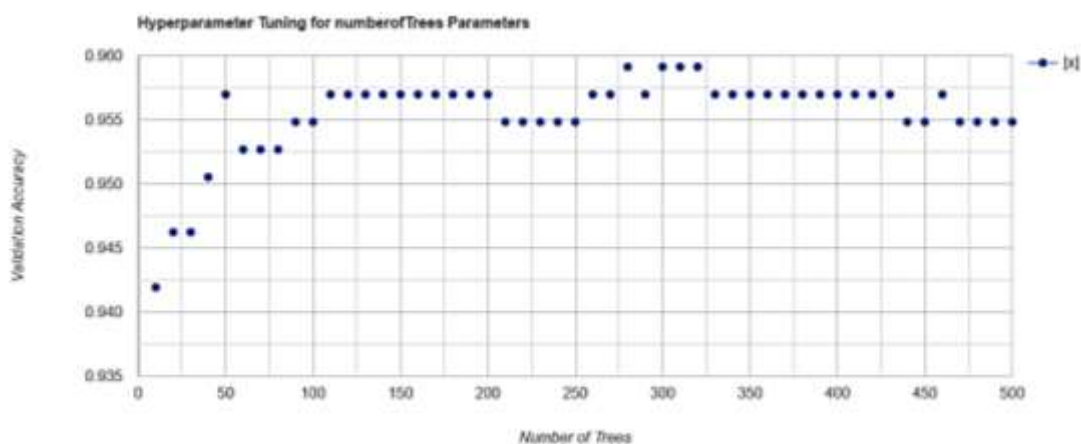
۳- نتایج و بحث

مطالعه حاضر باهدف طراحی روشی با دقت بالا برای طبقه‌بندی کاربری اراضی و مطالعه توسعه شهری تبریز در راستای بررسی افزایش سطوح نفوذناپذیر، انجام شد. همان‌طور که در بخش قبلی ذکر شد، روش‌شناسی از چهار بخش اصلی تشکیل شده است. ابتدا تصاویر لندست را معرفی گردید و داده‌های مرجع جهت آموزش ماشین و صحت‌سنجی الگوریتم جمع‌آوری شد. در مرحله بعد، باندهای طیفی استاندارد و شاخص‌های موردنظر را طبق شکل ۳ جهت طبقه‌بندی محاسبه و آنالیز شدند.

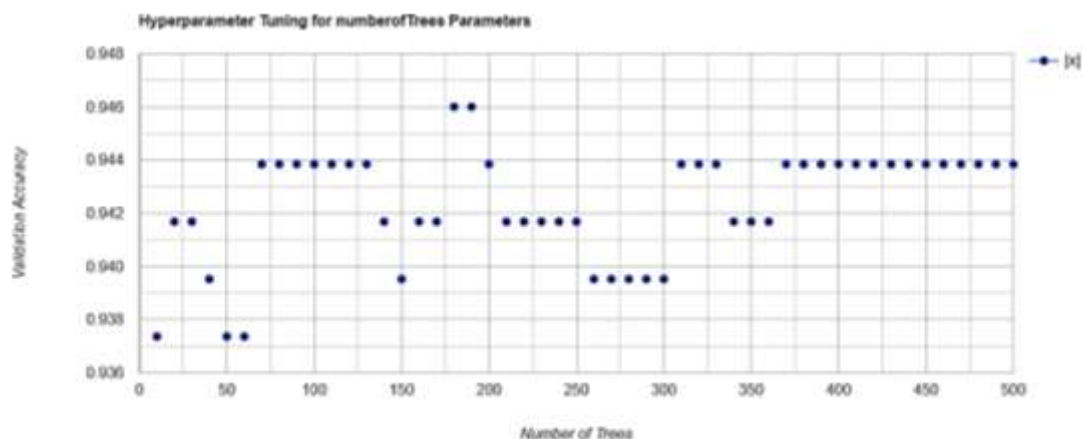


شکل ۳. تصاویر شاخص‌های طیفی NDVI, NDBI, MSAVI و UAI منطقه مورد مطالعه

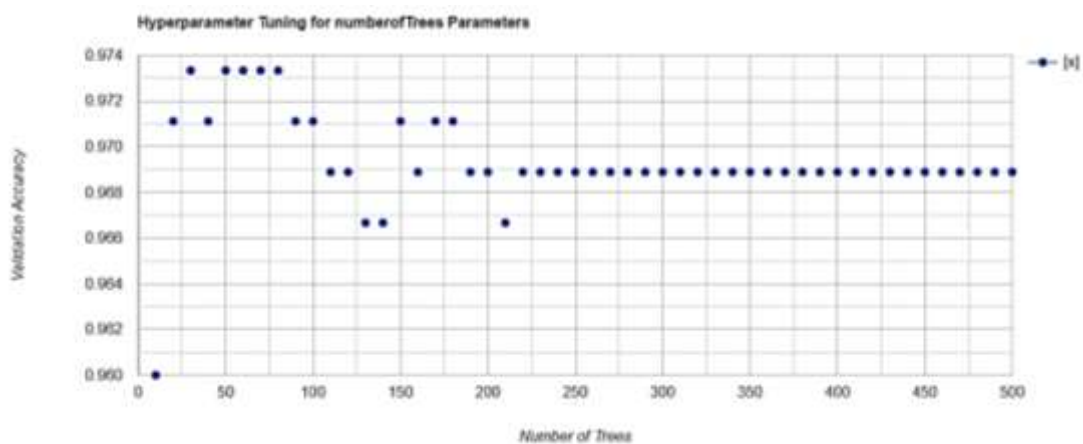
سپس از سه الگوریتم طبقه‌بندی نظارت شده KNN، SVM و RF برای طبقه‌بندی استفاده شد و بر اساس نتایج، الگوریتم RF نسبت به روش‌های دیگر از دقت بالاتری برخوردار هست؛ بنابراین الگوریتم RF برای طبقه‌بندی تصاویر انتخاب شد. جهت تنظیم هایپر پارامتر مدل جنگل تصادفی، از پارامتر تعداد درخت استفاده شد و مطابق نمودار شکل‌های ۴، ۵ و ۶ مقادیر هایپر پارامتر برای ۳ سال ۱۳۹۲، ۱۳۹۷ و ۱۴۰۲ که بالاترین دقت را در مجموعه داده آزمایشی به همراه داشت، به ترتیب ۲۸۰، ۱۸۰ و ۸۰ انتخاب شدند. تنظیم هایپر پارامتر یک مرحله مهم در آموزش مدل یادگیری ماشین است که امکان بهبود کیفیت مدل و دستیابی به نتایج بهتر در داده‌های آزمایشی را فراهم می‌کند. استفاده از تعداد زیاد درخت، بدلیل پیچیده تر شدن محاسبات، مدت زمان انجام آن افزایش می‌یابد، اما نتایج حاصل از دقت بالایی برخوردار خواهد شد و حتی می‌تواند منجر به کاهش دقت گردد (Le et al., 2022).



شکل ۴. نمودار هایپر پارامتر تیونینگ برای سال ۱۳۹۲



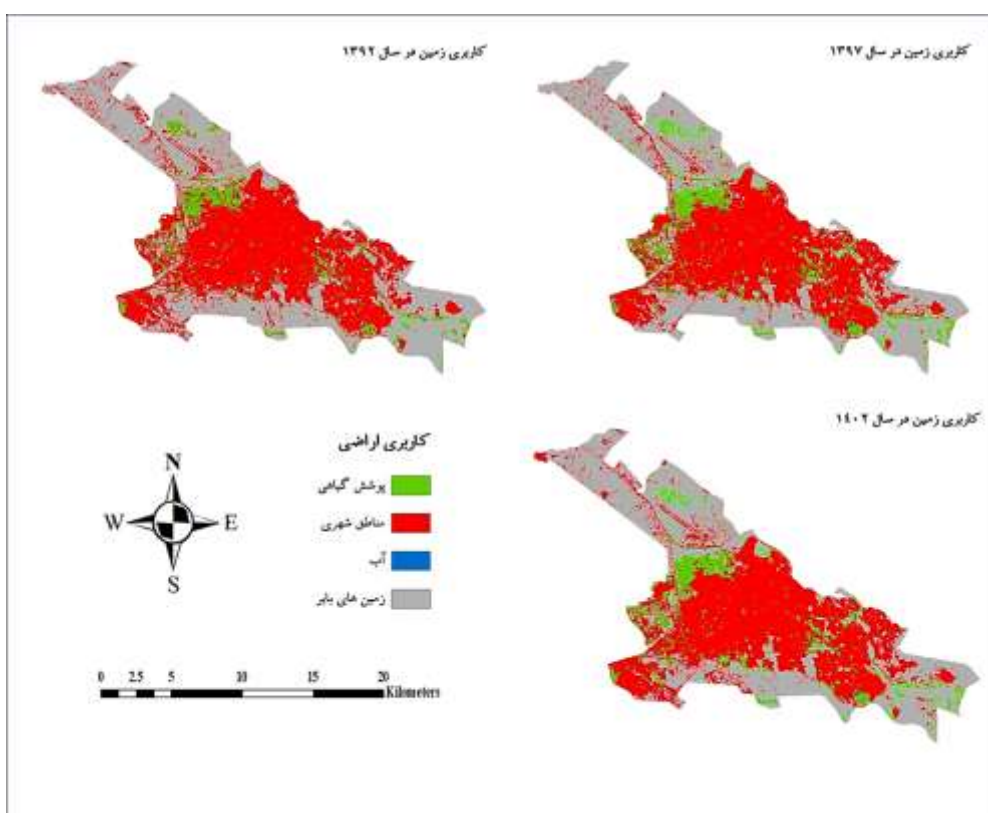
شکل ۵. نمودار هایپرپارامتر تیونینگ برای سال ۱۳۹۷



شکل ۶. نمودار هایپرپارامتر تیونینگ برای سال ۱۴۰۲

در نهایت طبقه‌بندی برای سه دوره زمانی انجام گردید. نقشه‌های کاربری زمین حاصل از نتایج مستخرج شده برای سه سال مورد مطالعه در شکل ۷ آورده شده است. برای ارزیابی دقت مدل طبقه‌بندی شده از دقت کلی (OA)، دقت تولیدکننده (PA)، دقت کاربر (UA)، امتیاز F1 و ضریب K استفاده گردید که در جدول ۳ و نمودارهای شکل ۸ مقادیر نشان داده شده اند. به طور کلی، تمام مجموعه داده ها دقت بالایی تولید کردند. دقت کلی برای سه سال به ترتیب برابر ۰,۹۵، ۰,۹۴، ۰,۹۷ و ضریب کاپا ۰,۹۴، ۰,۹۲، ۰,۹۶ بدست آمد. به طور

میانگین، مناطق شهری، بایر و پوشش گیاهی با بیشترین دقت طبقه‌بندی شدند. کمترین دقت در کلاس آب مشاهده شد.



شکل ۷. نتایج طبقه‌بندی پوشش کاربری اراضی برای سال‌های ۱۳۹۲، ۱۳۹۷ و ۱۴۰۲

جدول ۳. اندازه‌گیری دقت تولیدکننده و مصرف‌کننده، امتیاز F1، صحت کلی، و ضریب کاپا مربوط به هر کاربری برای طبقه‌بندی کاربری اراضی

سال	۱۳۹۲	۱۳۹۷	۱۴۰۲
کاربری اراضی	دقت کاربر	دقت کاربر	دقت کاربر
پوشش گیاهی	دقت تولیدکننده	دقت تولیدکننده	دقت تولیدکننده
	امتیاز F1	امتیاز F1	امتیاز F1
	۰/۹۵	۰/۹۴	۰/۹۷
	۰/۹۸	۰/۹۶	۰/۹۸
	۰/۹۷	۰/۹۵	۰/۹۸
	۰/۹۵	۰/۹۴	۰/۹۷
	۰/۹۸	۰/۹۶	۰/۹۸

۰/۹۶	۰/۹۵	۰/۹۶	۰/۹۲	۰/۹۵	۰/۹۱	۰/۹۵	۰/۹۶	۰/۹۴	شهری
۰/۹۶	۱	۰/۹۲	۰/۸۶	۰/۹۲	۰/۸	۰/۹	۰/۸۲	۱	آب
۰/۹۷	۰/۹۸	۰/۹۶	۰/۹۶	۰/۹۵	۰/۹۸	۰/۹۶	۰/۹۷	۰/۹۵	بایر
	۰/۹۷			۰/۹۴			۰/۹۵		صحت کلی
	۰/۹۶			۰/۹۲			۰/۹۴		ضریب کاپا



شکل ۸. نمودارهای دقت تولیدکننده و مصرف کننده، امتیاز F1 مربوط به هر کاربری

در سطح طبقاتی و براساس جدول ۴، نتایج حاکی از آن است که از سال ۱۳۹۲ تا سال ۱۴۰۲، مساحت شهری، پوشش گیاهی و آب افزایش یافته است. در مقابل، مناطق بایر روند کاهشی داشت. بر اساس مقایسه نقشه‌های کاربری، زمین‌های بایر عمدتاً به مناطق شهری تبدیل شدند. نتایج مطالعه برای ده سال اخیر، رشد شهری را بین سال‌های ۱۳۹۲ و ۱۴۰۲ نشان داد.

جدول ۴. تغییرات کاربری‌های مختلف در طول سه بازه زمانی

کاربری اراضی	۱۳۹۲ مساحت (km ²)	۱۳۹۷ مساحت (km ²)	۱۴۰۲ مساحت (km ²)
پوشش گیاهی	۱۵/۱	۲۱/۷۱	۲۰/۱۵
شهری	۱۰۶/۲۷	۱۰۳/۱۸	۱۰۷/۴۵
آب	۰/۰۵	۰/۰۷	۰/۰۷
بایر	۱۰۰/۵۴	۹۶/۹	۹۴/۲۱

بر اساس بررسی‌های صورت گرفته، نقشه‌های جهانی کاربری اراضی برای نگاشت دقیق نقشه‌های تغییر کاربری زمین نامناسب هستند و معرفی الگوریتم‌های جدید و داده‌های آموزشی در مناطق محلی ضروری است. تجزیه و تحلیل تغییرات نشان داد که تغییر در ساختارهای شهری از سال ۱۳۹۲ تا ۱۳۹۷ به صورت کاهشی بود که بر اساس مطالعات انجام گرفته ناشی از تخریب بافت فرسوده، تملک اراضی و آماده سازی جهت ساخت راه های دسترسی جدید و شهرک های جدیدالاحداث صنعتی و مسکونی بوده است و برای سال های ۱۳۹۷ تا ۱۴۰۲ گسترش مناطق شهری مشاهده شد.

۴- نتیجه گیری

این مطالعه یک روش توسعه یافته برای نقشه برداری کاربری اراضی با دقت بالا در مقیاس ۱۵ متر برای لندست ۸ بر اساس پلت فرم مبتنی برابر گوگل ارث انجین ارائه می کند. برای این منظور از الگوریتم جنگل تصادفی که یکی از قوی ترین طبقه بندی کننده ها بشمار می رود، استفاده شد. در این تحقیق تعداد متفاوتی از درختان بهینه بر اساس روش های پیرامتر تیونینگ بررسی گردید. سپس از نقشه CGLCL و تصاویر گوگل ارث با وضوح بالا برای ارائه برچسب های کلاس آموزشی قابل اعتماد استفاده شد. مقادیر دقت کلی، ضریب کاپا در بازه قابل قبولی قرار گرفتند. رویکردهای ارزیابی مقایسه ای نتایج طبقه بندی نهایی را تأیید کردند. نتایج به وضوح نشان داد که

و توسعه پایدار استفاده کرد. در این مطالعه تأثیر روش‌های پان تیز کردن مورد بررسی قرار نگرفت که برای مطالعات آینده پیشنهاد می‌گردد. جهت مطالعات آینده، اثربخشی مجموعه داده‌هایی مانند لندست ۹ و ترکیب لندست و سنتینل برای تغییر کاربری قابل بررسی است. برای بهبود دقت تجزیه و تحلیل تغییر کاربری نیز می‌توان ویژگی‌های بیشتری مانند مدل رقومی ارتفاعی زمین و شیب زمین اضافه کرد. همچنین می‌توان نتایج را با نقشه‌های حاصل از نقشه‌برداری میدانی مقایسه کرد.

۵-منابع

مقالات انگلیسی

- Amani, M., Ghorbanian, A., Ahmadi, S. A., Kakooei, M., Moghimi, A., Mirmazloumi, S. M., Parsian, S. (2020). Google earth engine cloud computing platform for remote sensing big data applications: A comprehensive review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 5326-5350 .
- Aryal, J., Sitaula, C., & Frery, A. C. (2023). Land use and land cover (LULC) performance modeling using machine learning algorithms: a case study of the city of Melbourne, Australia. *Scientific Reports*, 13(1), 13510 .
- Avand, M., & Moradi, H. (2021). Using machine learning models, remote sensing, and GIS to investigate the effects of changing climates and land uses on flood probability. *Journal of Hydrology*, 595, 125663 .
- Bai, L., Zhou, X., Tian, Y., & Wei, L. (2024). Evaluation for the Development of Urban Agglomeration Integration: A Case Study of Pearl River Delta. *Social Indicators Research*, 1-28 .
- Belgiu, M., & Drăguț, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS journal of photogrammetry and remote sensing*, 114, 24-31 .
- Chapa, F., Hariharan, S., & Hack, J. (2019). A new approach to high-resolution urban land use classification using open access software and true color satellite images. *Sustainability*, 11(19), 5266 .

- Deng, J., Pan, S., Zhou, M., Gao, W., Yan, Y., Niu, Z., & Han, W. (2023). Optimum sampling window size and vegetation index selection for low-altitude multispectral estimation of root soil moisture content for Xuxiang Kiwifruit. *Agricultural Water Management*, 282, 108297 .
- Egorov, A., Hansen, M., Roy, D. P., Kommareddy, A., & Potapov, P. (2015). Image interpretation-guided supervised classification using nested segmentation. *Remote Sensing of Environment*, 165, 135-147 .
- Foody, G. M. (2020). Explaining the unsuitability of the kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification. *Remote sensing of environment*, 239, 111630 .
- Ghiyasi, B., Noori, R., Sheikhan ,H., Zeynolabedin, A., Sun, Y., Jun, C., . . . Abolfathi, S. (2022). Uncertainty quantification of granular computing-neural network model for prediction of pollutant longitudinal dispersion coefficient in aquatic streams. *Scientific reports*, 12(1), 4610 .
- Gorji, T., Yildirim, A., Hamzhepour, N., Tanik, A., & Sertel, E. (2020). Soil salinity analysis of Urmia Lake Basin using Landsat-8 OLI and Sentinel-2A based spectral indices and electrical conductivity measurements. *Ecological Indicators*, 112, 106173 .
- Guha, S., Govil, H., Gill, N., & Dey, A. (2021). A long-term seasonal analysis on the relationship between LST and NDBI using Landsat data. *Quaternary International*, 575, 249-258 .
- Hashemi Aslani, Z., Omidvar, B., & Karbassi, A. (2022). Integrated model for land-use transformation analysis based on multi-layer perception neural network and agent-based model. *Environmental Science and Pollution Research*, 29(39), 59770-59783 .
- Heydari, S. S., & Mountrakis, G. (2018). Effect of classifier selection, reference sample size, reference class distribution and scene heterogeneity in per-pixel classification accuracy using 26 Landsat sites. *Remote Sensing of Environment*, 204, 648-658 .
- Huang, S., Tang, L., Hupy, J. P., Wang, Y., & Shao, G. (2021). A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. *Journal of Forestry Research*, 32(1), 1-6 .
- Karimi, F., Sultana, S., Babakan, A. S., & Suthaharan, S. (2019). An enhanced support vector machine model for urban expansion prediction. *Computers, Environment and Urban Systems*, 75, 61-75 .

Le, T. D. H., Pham, L. H., Dinh, Q. T., Hang, N. T. T., & Tran, T. A. T. (2022). Rapid method for yearly LULC classification using Random Forest and incorporating time-series NDVI and topography: a case study of Thanh Hoa province, Vietnam. *Geocarto International*, 37(27), 17200-17215 .

Li, L., Yu, Q., Gao, L., Yu, B., & Lu, Z. (2021). The effect of urban land-use change on runoff water quality: a case study in Hangzhou City. *International Journal of Environmental Research and Public Health*, 18(20), 10748 .

Li, S., Xu, L., Jing, Y., Yin, H., Li, X., & Guan, X. (2021). High-quality vegetation index product generation: A review of NDVI time series reconstruction techniques. *International Journal of Applied Earth Observation and Geoinformation*, 105, 102640 .

Palsson, F., Sveinsson, J. R., Benediktsson, J. A., & Aanaes, H. (2011). Classification of pansharpened urban satellite images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(1), 281-297 .

Ren, Y., Lü, Y., Comber, A., Fu, B., Harris, P., & Wu, L. (2019). Spatially explicit simulation of land use/land cover changes: Current coverage and future prospects. *Earth-Science Reviews*, 190, 398-41 .^o

Sajikumar, N., & Remya, R. (2015). Impact of land cover and land use change on runoff characteristics. *Journal of environmental management*, 161, 460-468 .

Salcedo-Sanz, S., Rojo-Álvarez, J. L., Martínez-Ramón, M., & Camps-Valls, G. (2014). Support vector machines in engineering: an overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 4(3), 234-267 .

Speiser, J. L., Miller, M. E., Tooze, J., & Ip, E. (2019). A comparison of random forest variable selection methods for classification prediction modeling. *Expert systems with applications*, 134, 93-101 .

Talukdar, S., Singha, P., Mahato, S., Pal, S., Liou, Y.-A., & Rahman, A. (2020). Land-use land-cover classification by machine learning classifiers for satellite observations—A review. *Remote Sensing*, 12(7), 1135 .

Tamiminia, H., Salehi, B., Mahdianpari, M., Quackenbush, L., Adeli, S., & Brisco, B. (2020). Google Earth Engine for geo-big data applications: A meta-analysis and systematic review. *ISPRS journal of photogrammetry and remote sensing*, 164, 152-170 .

Tang, J., Liu, D., Shang, C., & Niu, J. (2024). Impacts of land use change on surface infiltration capacity and urban flood risk in a representative karst mountain city over the last two decades. *Journal of Cleaner Production*, 454, 142196 .

Tassi, A., & Vizzari, M. (2020). Object-oriented lule classification in google earth engine combining snic, glcm, and machine learning algorithms. *Remote Sensing*, 12(22), 3776 .

Vafa, H., & Seker, D. Z. (2023). Flood risk susceptibility and evaluation by AHP technique in Denizli, Turkiye. *Intercontinental Geoinformation Days*, 6, 53-56 .

Weber, A., Fohrer, N., & Möller, D. (2001). Long-term land use changes in a mesoscale watershed due to socio-economic factors—effects on landscape structures and functions. *Ecological modelling*, 140(1-2), 125-140 .

Yang, L., Driscoll, J., Sarigai, S., Wu, Q., Chen, H., & Lippitt, C. D. (2022). Google Earth Engine and artificial intelligence (AI): a comprehensive review. *Remote Sensing*, 14(14), 3253 .

Yu, L., Su, J., Li, C., Wang, L., Luo, Z., & Yan, B. (2018). Improvement of moderate resolution land use and land cover classification by introducing adjacent region features. *Remote Sensing*, 10(3), 414 .

Zheng, Y., Tang, L., & Wang, H. (2021). An improved approach for monitoring urban built-up areas by combining NPP-VIIRS nighttime light, NDVI, NDWI, and NDBI. *Journal of Cleaner Production*, 328, 129488 .

مقالات فارسی

Alavi S.A., Raushit S., Yousefi M., and Kia, R. (2016). Revealing urban land use changes using satellite image processing based on neural network (case study: Tabriz city). *Geographical Space*, 17(59), 27-46. (In Persian)

al-Madrasi al-Hosseini S.A., Kerami, J., and Roshan Bakhsh, S. (2013). Investigating land use changes in Hamadan between 2002-2009 using artificial neural network (MLP) and LCM,

National Conference on the Application of Advanced Spatial Analysis Models (Remote Sensing and GIS) in Land Management, Yazd. (In Persian)

Banan Ferdowsi F., and Din Pajouh, Y. (2018). Investigating the temporal distribution of daily rainfall with the method of normal rainfall curves (NRC) (case study: Tabriz station). *Water and Soil Science (Agricultural Science)*, 29(2), 1-14. (In Persian)

Jafari F, Mozni M, and Badli, A. (2019). Future research of urban land use changes in Tabriz metropolis. *Spatial Planning*, 10(2(37)), 1-22. (In Persian)

Mousavi, M., and Yazdani Chaharbarj, R. (2014). Analyzing suitability of land use for Tabriz city development using AHP-OWA model. *Urban Planning Geography Research*, 3(3), 361-381. (In Persian)



Assessment and Analysis of Urban Development and Land Use Change Impact on Surface Permeability Using Machine Learning Algorithm and Remote Sensing Data: A Case Study of Tabriz City

Abstract

Optimal classification of satellite images enhances the extraction of spatial-temporal information for land use and land cover classification, enabling the examination of various socio-economic and environmental impacts. Initially, three supervised classification algorithms—K-Nearest Neighbors, Support Vector Machines, and Random Forests—were utilized with default parameters. The results indicated that the Random Forest algorithm outperformed the others in accuracy. Consequently, the Random Forest classification algorithm on the Google Earth Engine platform, using Landsat 8 images as input for 2013 to 2023, was employed. The impact of different image combinations based on several spectral indices on the final classification accuracy was assessed. The findings show that using Landsat products (TOA) and the GEE platform provides faster and more accurate classification, with optimal hyperparameter values for 2013, 2018, and 2023 being 280, 180, and 80, respectively. The findings indicate an increase in vegetation cover alongside gradual urban expansion, reflecting stable development policies. However, this growth is primarily attributed to the establishment of large leisure areas and tree planting on the city outskirts. Despite initiatives aimed at reducing air pollution and promoting community well-being, vegetation in central urban areas has either remained stagnant or declined in certain areas. Considering the implications of climate change and increased heavy rainfall, the lack of vegetation in central zones may lead to the formation of impermeable surfaces, exacerbating the risk of flash floods. This underscores the critical need for enhanced research focus and proactive measures in this field.

Keywords: Land use, Google Earth Engine, Hyperparameter tuning, Random Forest, Landsat