



سنجش از دور

GIS ایران



سنجش از دور و GIS ایران
Iranian Remote Sensing & GIS

سال سیزدهم، شماره اول، بهار ۱۴۰۰
Vol.13, No. 1, Spring 2021

۱۷-۲۲

مقاله پژوهشی

استخراج تاج پوشش درختان شهری با روش طبقه بندی شیء پایه و الگوریتم های یادگیری ماشین

ناهید حق شناس^۱، علی شمس الدینی*^۲، حسین عقیقی^۳

۱. کارشناس ارشد سنجش از دور و سیستم اطلاعات جغرافیایی، دانشگاه تربیت مدرس

۲. استادیار گروه سنجش از دور و سیستم اطلاعات جغرافیایی، دانشگاه تربیت مدرس، تهران

۳. استادیار گروه سنجش از دور و سیستم اطلاعات جغرافیایی، دانشگاه شهید بهشتی، تهران

تاریخ پذیرش مقاله: ۱۳۹۹/۰۶/۱۰

تاریخ دریافت مقاله: ۱۳۹۸/۰۹/۱۰

چکیده

آگاهی از میزان تاج پوشش درختان در مناطق شهری به علت تأثیرات آن در کاهش آلودگی های هوا، کاهش آلودگی صوتی، جلوگیری از باد، ذخیره نزولات آسمانی، و کاهش و کنترل رواناب های شهری بسیار ضروری است. از آن جاکه استخراج تاج پوشش درختان با روش های دستی بسیار وقت گیر و پرهزینه است، تکنیک های سنجش از دور می توانند ابزار مناسبی برای تأمین این داده ها باشند. در طبقه بندی شیء پایه، انتخاب پارامترهای بهینه قطعه بندی، به ویژه پارامتر مقیاس، اهمیت بسیاری دارد و معمولاً با شیوه آزمایش و خطا تعیین می شود که کاملاً تجربی است. بنابراین، یکی از اهداف این پژوهش انتخاب مقیاس بهینه قطعه بندی به صورت خودکار است. همچنین، پس از استخراج قطعات، لازم است با یک روش طبقه بندی، قطعات استخراج شده تعیین کاربری/ پوشش زمین شوند و در این زمینه، انتخاب نوع روش طبقه بندی در نتیجه نهایی طبقه بندی شیء پایه بسیار اهمیت دارد. از این رو، پس از طبقه بندی با استفاده از داده های لیدار و تصاویر هوایی از شهر واهینگن در آلمان و تعیین ویژگی های مهم مستخرج از قطعات، با استفاده از روش انتخاب ویژگی بر مبنای جنگل تصادفی، قطعات مربوط به تاج پوشش درختان از سایر قطعات تفکیک شد. این کار با بهره گرفتن از شیوه های یادگیری ماشین شامل ماشین بردار پشتیبان، جنگل تصادفی و درخت تصمیم گیری صورت گرفت. نتایج نشان دهنده برتری الگوریتم ماشین بردار پشتیبان، به منزله برترین الگوریتم طبقه بندی کننده، و مقیاس ۲۵، به منزله بهترین مقیاس انتخابی، بود و در نهایت، الگوریتم های ماشین بردار پشتیبان، جنگل تصادفی و درخت تصمیم گیری در مقیاس ۲۵، به ترتیب، با شاخص های کیفیت ۷۹.۹۰ و ۷۹.۱۶ و ۷۶.۹۰ توانستند تاج پوشش درختان را استخراج کنند.

کلیدواژه ها: تاج پوشش درختان شهری، روش شیء پایه، پارامتر مقیاس، الگوریتم های یادگیری ماشین.

* نویسنده مکاتبه کننده: دانشگاه تربیت مدرس، دانشکده علوم انسانی، گروه سنجش از دور و سیستم اطلاعات جغرافیایی، تلفن: ۰۹۱۵۵۰۳۴۵۲۲

۱- مقدمه

در فضاهای سبز شهری، تاج‌پوشش^۱ درختان مهم‌ترین بخش این فضاهاست که نقش بسیار مهمی در کاهش آلودگی هوا، کاهش آلودگی صوتی، جلوگیری از باد، کاهش جزایر گرمایی شهری، و کاهش و کنترل رواناب‌های شهری دارد و می‌توان آن را به‌منزلهٔ ابزاری مؤثر در کنترل کیفیت زندگی شهری و کاهش خسارات وارد شده به این مکان‌ها دخیل دانست. بنابراین، آگاهی از میزان و پراکنش تاج‌پوشش درختان در مناطق شهری اهمیت بسیاری در زمینهٔ مدیریت شهری دارد (Mathieu & Aryal, 2005). از این‌رو، با آگاهی از مساحت و نحوهٔ پراکنش آنها، می‌توان به‌حداقل‌رساندن تأثیرات زیان‌بار شهرنشینی را به‌درستی برنامه‌ریزی و مدیریت کرد. اطلاعات درمورد تاج‌پوشش درختان از طریق آماربرداری‌های زمینی گردآوری می‌شود اما روش‌های متفاوت آماربرداری زمینی دارای دقت متفاوت‌اند و وقت‌گیر و هزینه‌برند. بنابراین، پژوهشگران استفاده از داده‌های سنجنش از دور را برای استخراج تاج‌پوشش درختان و مدیریت فضاهای سبز شهری پیشنهاد کرده‌اند (Parmehr et al., 2016). طبقه‌بندی یکی از پرکاربردترین روش‌های استخراج اطلاعات در سنجنش از دور محسوب می‌شود. در روش‌های معمول طبقه‌بندی، واحد پردازش بر مبنای پیکسل استوار است که در آن از خصوصیات طیفی برای استخراج ویژگی‌ها استفاده می‌شود؛ در حالی که در طبقه‌بندی شیء‌پایه^۲، علاوه بر بررسی ویژگی طیفی پیکسل‌ها، مجموعه‌ای از ویژگی‌های طیفی و بافتی و شکل‌قطعاتی، شامل گروهی از پیکسل‌های مشابه، مورد توجه قرار می‌گیرد. در طبقه‌بندی پیکسل‌مبنا^۳، با افزایش واریانس طیفی در کلاس‌ها، جداسازی عوارض دشوارتر می‌شود. بنابراین، طبقه‌بندی تصاویر با قدرت تفکیک مکانی بالا، با استفاده از روش‌های پیکسل‌مبنا، نتایج ضعیفی در پی خواهد داشت (Shaban & Dikshit, 2001). چن و همکاران^۴ (۲۰۰۹)، با استفاده از طبقه‌بندی شیء‌پایه و

پیکسل‌پایه، مناطق شهری را طبقه‌بندی کردند و نتایج پژوهش آنها افزایش ۲۸.۳ درصدی صحت کلی طبقه‌بندی را در روش شیء‌پایه، به‌نسبت روش پیکسل‌مبنا، نشان داد. ریچار رحمان^۵ (۲۰۰۸) با استفاده از داده‌های IRS اقدام به طبقه‌بندی و تهیهٔ نقشهٔ پوشش زمین کرد. با تغییر پارامتر مقیاس در سه سطح متفاوت (۲۰، ۲۵ و ۳۰)، نتایج نشان‌دهندهٔ برتری روش شیء‌پایه، با دقت ۹۲٪، بر روش پیکسل‌مبنا، با دقت ۸۶٪، بوده است. همچنین، نتایج نشان داد که طبقه‌بندی شیء‌پایه باعث ایجاد تمایزی پایه بین کلاس‌های متفاوت می‌شود. در روش طبقه‌بندی شیء‌پایه، برای طبقه‌بندی، در مرحلهٔ نخست باید پیکسل‌ها بر مبنای شباهت طیفی و بافتی قطعه‌بندی^۶ شوند و پس از آن، برای طبقه‌بندی قطعات ایجاد شده اقدام شود (Baatz, 2000)؛ بنابراین، قطعه‌بندی تصویر مرحله‌ای بسیار مهم و تأثیرگذار در این شیوه به‌شمار می‌رود (Agarwal et al., 2013). روش‌های گوناگونی به‌منظور قطعه‌بندی تصویر توسعه یافته است. یکی از این روش‌های پرکاربرد الگوریتم قطعه‌بندی چندمقیاسه^۷ است. در این الگوریتم، مقیاس نقش بسیار مهمی در تعیین قطعات ایفا می‌کند؛ به طوری که ویژگی‌های قطعات تصویری مانند بافت، طیف و شکل تاحدودی با توجه به انتخاب مقیاس مشخص می‌شود (Baatz, 2000).

استخراج سطح تاج‌پوشش درختان شهری یکی از مباحث مهم در طبقه‌بندی کاربری/پوشش مناطق شهری است. موسا و الشیمی^۸ (۲۰۱۲) در ناحیهٔ شهری آلمان، با استفاده از ترکیب تصویر هوایی و لیدار، به استخراج ساختمان‌ها و درختان به‌روش شیء‌پایه پرداختند.

1. Tree Crown
2. Object Base
3. Pixel Base
4. Chen
5. Rejaur Rahman
6. Segmentation
7. Multi Resolutions Segmentation
8. Moussa & El Sheimy

مناطق گوناگون، مقایسه‌ای انجام دادند. نتایج آنها نشان داد شبکه عصبی می‌تواند، در مقایسه با سایر الگوریتم‌ها، عملکرد بهتری داشته باشد (Zhang et al., 2018). تحقیقات گذشته نشان می‌دهد، به‌رغم اینکه مطالعات بسیاری درباره طبقه بندی شیء پایه و پیکسل پایه صورت گرفته است؛ در بیشتر موارد، استخراج کلاس درختان با صحتی کمتر از دیگر کلاس‌ها انجام شده و به بحث قطعه بندی و چگونگی انتخاب مقیاس بهینه در جریان قطعه بندی در روش شیء پایه کمتر توجه شده است. به این دلیل که این مرحله، در طبقه بندی شیء پایه، بسیار مهم است و می‌تواند صحت و سرعت طبقه بندی را تا حدود زیادی تحت تأثیر قرار دهد؛ در تحقیق پیش رو، با تمرکز بر طبقه بندی شیء پایه و انتخاب مقیاس مناسب قطعه بندی با روشی خودکار، به افزایش صحت استخراج تاج پوشش درختان پرداخته می‌شود. از دیگر اهداف این تحقیق مقایسه روش‌های یادگیری ماشین شامل روش ماشین بردار پشتیبان، روش جنگل تصادفی و روش درخت تصمیم‌گیری، به منظور تعیین کاربری/ پوشش قطعات تفکیک شده، است.

۲- منطقه مورد مطالعه و داده‌های مورد استفاده

منطقه مورد مطالعه شهر واهینگن^{۱۱}، واقع در ایالت بادن وورتمبرگ^{۱۲}، در جنوب کشور آلمان است. این منطقه، با مساحت ۵۴۴۱۸ مترمربع، در طول

شاخص کیفیت برای استخراج درختان، در سه منطقه متفاوت، به ترتیب، ۳۴، ۵۷ و ۵۲٪ و شاخص کیفیت نیز، برای استخراج ساختمان‌ها، به ترتیب، ۸۴ و ۸۹ و ۸۳٪ بود. نیمیر^۱ و همکاران (۲۰۱۳)، با استفاده از روش جنگل تصادفی شرطی^۲ و داده‌های نقطه‌ای لیدار، مناطق شهری را در سه ناحیه شهری آلمان طبقه بندی کردند. آنها، در مجموع، از ۷۱ ویژگی داده‌های لیدار بهره بردند. در این تحقیق، ساختمان‌ها به خوبی تشخیص داده شد اما درختان با صحت بالایی تشخیص داده نشدند. در تحقیقی دیگر، مکفادن^۳ و همکاران (۲۰۱۲)، با طبقه بندی شیء پایه و استفاده ترکیبی از لایه‌های سنجش از دور شامل مدل رقومی نرمال شده سطح^۴، انحراف استاندارد نقاط ارتفاعی لیدار، شدت، تصاویر چندطیفی و نیز با استفاده از لایه‌های موضوعی، کلاس‌های متفاوت را در مناطق شهری طبقه بندی کردند. در تحقیق آنها، لایه‌های مدل رقومی نرمال شده سطح و انحراف استاندارد از ارتفاع داده‌های لیدار، به همراه لایه شدت، برای جداسازی درختان از سایر کلاس‌ها بسیار مهم تشخیص داده شد. گرک و خیاثو^۵ (۲۰۱۴) نیز، با تلفیق ابرنقاط لیدار و تصویر هوایی با قدرت تفکیک مکانی زیاد، از طبقه بندی کننده‌های نظارت شده^۶ و نظارت نشده^۷ و داده‌های ارتفاعی و طیفی در طبقه بندی مناطق شهری استفاده کردند. نتایج آنها بیانگر تأثیرگذاری مطلوب شاخص NDVI در جداسازی مناطق گیاهی از غیرگیاهی، به جز در مکان‌های سایه، بود.

همان گونه که بیان کردیم، یکی از مهم‌ترین مراحل طبقه بندی شیء پایه استفاده از الگوریتم‌های طبقه بندی به منظور تعیین نوع کاربری/ پوشش قطعات تفکیک شده است و انتخاب روش طبقه بندی مناسب، بدین منظور، می‌تواند در تعیین صحت نهایی طبقه بندی مؤثر واقع شود. در یکی از معدود تحقیقاتی که در زمینه مقایسه روش‌های طبقه بندی با هدف تعیین نوع کاربری/ پوشش قطعات انجام شده است، سارگنت^۸ و همکاران (۲۰۱۸) بین الگوریتم‌های طبقه بندی ماشین بردار پشتیبان^۹ و شبکه عصبی پرسپترون چندلایه^{۱۰}، برای

1. Niemeyer
2. Conditional Random Forest
3. MacFaden
4. Normalized Digital Surface Modeling
5. Xiao & Gerke
6. Supervized
7. Unsupervised
8. Sargent
9. Support Vector Machine
10. Multi-layer Perceptron Neural Network
11. Vaihingen
12. Baden-Wuerttemberg

۳-۱- پیش پردازش داده‌های لیدار

برای فیلتر کردن داده‌های لیدار و جداسازی نقاط زمینی از غیرزمینی، الگوریتم فیلتر شیب‌مبنا به کار رفت و لایه مدل رقومی زمین، با استفاده از نقاط زمینی انتخابی، از طریق الگوریتم شیب‌مبنا و روش درون‌یابی معکوس وزنی فاصله، ایجاد شد. لایه مدل رقومی سطح نیز با استفاده از نقاط برگشتی اول داده‌های لیدار ایجاد شد. سپس، با کم کردن لایه مدل رقومی زمین از مدل رقومی سطح، مدل رقومی نرمال‌شده سطح و از این لایه نیز، لایه شیب پدید آمد.

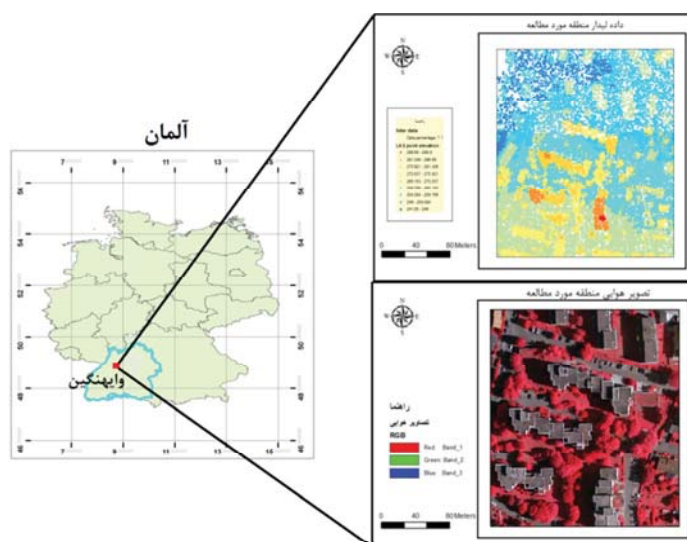
۳-۲- قطعه‌بندی

در این تحقیق، از سه باند اصلی تصویر هوایی و لایه مدل رقومی نرمال‌شده سطح، برای ورود به مرحله قطعه‌بندی، استفاده شد. با توجه به اینکه یکی از روش‌های پرکاربرد در قطعه‌بندی شیوه قطعه‌بندی چندمقیاسه است، در این مطالعه، این روش به کار رفت. این الگوریتم با یک عامل مقیاس و یک معیار ناهمگونی^۱ کنترل شد.

جغرافیایی ۸ دقیقه و ۵۷ درجه و ۲۵ ثانیه تا ۸ دقیقه و ۵۷ درجه و ۳۵ ثانیه شرقی و در عرض جغرافیایی ۴۸ دقیقه و ۵۵ درجه و ۳۷ ثانیه تا ۴۸ دقیقه و ۵۵ درجه و ۴۵ ثانیه شمالی قرار دارد. مجموعه داده‌های مورد استفاده در این تحقیق، تصاویر هوایی با قدرت تفکیک مکانی بالا و داده‌های نقطه‌ای لیدار تهیه‌شده از شهر واهینگن است. تصاویر تهیه‌شده با دوربین DSM Intergraph/ZI، با فاصله کانونی ۱۲۰ میلی‌متر، از ارتفاع ۹۰۰ متری و در سال ۲۰۰۶ برداشت شده است که توانایی برداشت سه باند طیفی مادون قرمز نزدیک، قرمز و سبز را با قدرت تفکیک مکانی ۹ سانتی‌متر دارد. داده‌های لیدار منطقه نیز با استفاده از سیستم Lecia ALS50 و در ده نوار، از ارتفاع ۵۰۰ متری، با میزان پوشش ۳۰٪، از شهر مورد نظر برداشت شده‌اند. میزان تراکم نقاط، در مناطقی که زیر پوشش دو نوار مجاور قرار گرفته‌اند، معادل با ۶۸ نقطه در هر مترمربع و در مناطق زیر پوشش با یک نوار برداشتی، به‌طور متوسط، ۴ نقطه در هر مترمربع است.

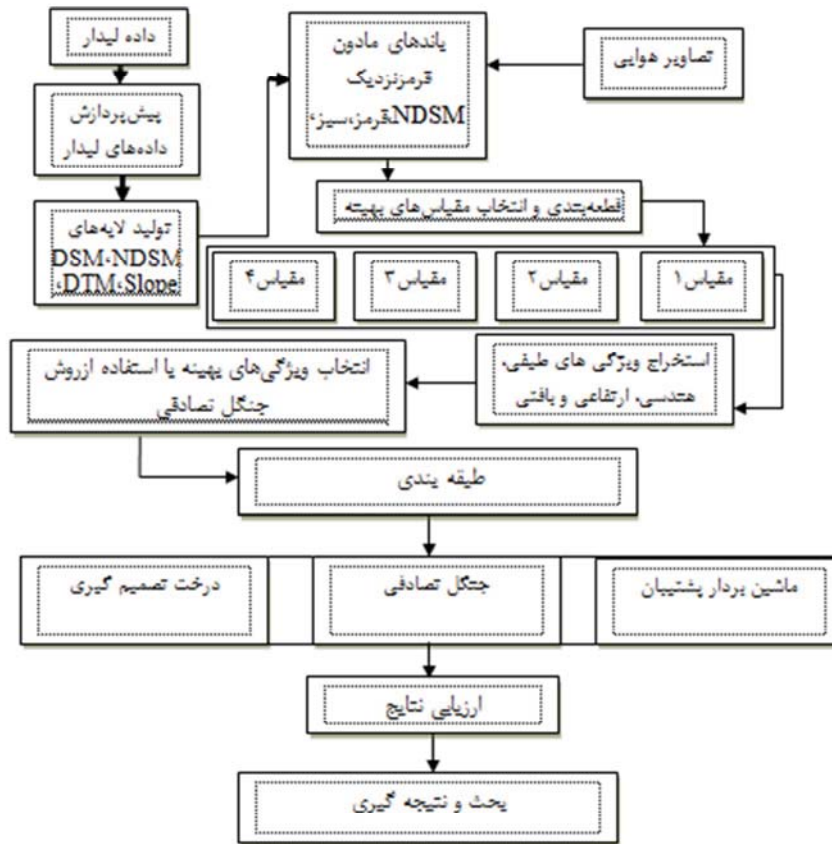
۳- روش تحقیق

نمودار جریان تحقیق، در شکل ۲، مراحل کار در این پژوهش را نشان می‌دهد.



شکل ۱. منطقه مورد مطالعه

1. Heterogeny



شکل ۲. نمودار جریان تحقیق

به دلیل نبود پارامترهای ایده آل برای استخراج عوارض گوناگون، پارامترهای دیگر با توجه به داده های مورد استفاده و عوارض استخراجی مورد نظر و عارضه های غالب در تصویر و با سعی و خطا، طوری که میزان Under segment و Over segment حداقل مقدار ممکن بود، تنظیم شد. بر این اساس، وزن پارامتر ناهمگنی شکلی ۰.۴، ناهمگنی طیفی ۰.۶، پارامترهای نرمی ۰.۵ و فشردگی ۰.۵ در نظر گرفته شد.

۳-۳- استخراج ویژگی

در این مرحله و پس از قطعه بندی، ویژگی های بافتی، طیفی، ارتفاعی و هندسی از قطعه های ایجاد شده استخراج شد. برای ویژگی های بافتی، ماتریس وقوع

به منظور انتخاب مقیاس بهینه در روش چندمقیاسه، از معیار واریانس محلی^۱ ناهمگونی قطعات استفاده شد. در این شیوه، برای نشان دادن نحوه رفتار واریانس محلی، زمان گذار از یک مقیاس به مقیاس دیگر، از معیار نرخ تغییر^۲ استفاده شده و برای محاسبه نرخ تغییر در این مطالعه نیز، رابطه (۱) به کار رفته است. در رابطه (۱)، L_{LV} واریانس محلی در مقیاس مورد نظر و $(L_{LV} - 1)$ واریانس محلی در مقیاس پایین تر را نشان می دهد. فرض بر این است که نقاط اوج، در نمودار دامنه تغییر واریانس محلی، مقیاسی را نشان می دهد که در آن می توان تصویر را در مناسب ترین وضعیت خود به نسبت ویژگی های داده، قطعه بندی کرد (Drăguț et al., 2010).

$$ROC = \left[\frac{L_{LV} - (L_{LV} - 1)}{L_{LV} - 1} \right] \cdot 100 \quad \text{رابطه (۱)}$$

1. Local Variance
2. Rate of Change

آموزشی استفاده شود و نیز، برای ارزیابی نتایج صحت طبقه‌بندی، بتوان صحت طبقه‌بندی را درمورد قطعات محاسبه کرد. شکل ۳ نحوه تقسیم داده‌های آموزشی و آزمایشی را نشان می‌دهد.

۳-۵- طبقه‌بندی

در این تحقیق، از الگوریتم‌های طبقه‌بندی درخت تصمیم‌گیری، جنگل تصادفی و ماشین بردار پشتیبان برای طبقه‌بندی و تهیه نقشه تاج‌پوشش درختان، با روش شیء‌پایه، استفاده شد.

۳-۵-۱- الگوریتم درخت تصمیم‌گیری

الگوریتم درخت تصمیم‌گیری یا درخت طبقه‌بندی^{۲۴} یکی از الگوریتم‌های طبقه‌بندی ناپارامتریک^{۲۵} است.

توأم گام‌های خاکستری به کار رفت که در مجموع، ۲۴ ویژگی بافتی (جدول ۱) برای باندهای مادون قرمز نزدیک، قرمز و سبز استخراج شد. از میانگین و انحراف معیار باندهای اصلی و شاخص‌های گیاهی NDVI، SAVI، RVI و DVI به‌منزله ویژگی‌های طیفی و از میانگین و انحراف معیار لایه‌های ارتفاعی حاصل از داده لیدار به‌منزله ویژگی‌های ارتفاعی استفاده شد. شانزده ویژگی نیز، با توجه به شکل قطعات ایجادشده در مرحله قطعه‌بندی (جدول ۱)، به‌صورت ویژگی‌های هندسی در نظر گرفته شد. در مرحله بعد، با توجه به ابعاد زیاد این ویژگی‌ها، الگوریتم جنگل تصادفی برای کاهش ابعاد ویژگی به کار رفت و از مجموع ۶۵ ویژگی، ۲۴ تا به‌منزله ویژگی‌های پراهمیت انتخاب شدند تا، در روش‌های یادگیری ماشین، برای طبقه‌بندی استفاده شوند.

جدول ۱. ویژگی‌های طیفی، هندسی، ارتفاعی و بافتی استخراج‌شده

| | |
|-------------------|--|
| ویژگی‌های طیفی | میانگین باندهای مادون قرمز نزدیک، قرمز، سبز، مؤلفه‌های اول، دوم و سوم حاصل از تحلیل مؤلفه‌های اصلی انحراف استاندارد باندهای مادون قرمز نزدیک، قرمز، سبز، مؤلفه‌های اول، دوم و سوم حاصل از تحلیل مؤلفه‌های اصلی شاخص‌های گیاهی NDVI، SAVI، RVI و DVI |
| ویژگی‌های بافتی | شامل تجانس ^۱ ، نداشتن تشابه ^۲ ، واریانس ^۳ ، بی‌نظمی ^۴ ، میانگین ^۵ ، همبستگی ^۶ ، همگنی ^۷ ، زاویه دوم لحظه‌ای ^۸ |
| ویژگی‌های ارتفاعی | میانگین و انحراف استاندارد لایه‌های مدل رقومی زمین، مدل رقومی نرمال‌شده سطح، مدل رقومی سطح و شیب |
| ویژگی‌های هندسی | شاخص مرز ^۹ ، نامتقارنی ^{۱۰} ، فشردگی ^{۱۱} ، چگالی ^{۱۲} ، انطباق بیضوی ^{۱۳} ، انطباق دایره‌ای ^{۱۴} ، شاخص شکل ^{۱۵} ، انطباق مستطیلی ^{۱۶} ، جهت ^{۱۷} ، طول بر عرض ^{۱۸} ، مساحت ^{۱۹} ، عرض ^{۲۰} ، شعاع کوچک‌ترین بیضی محصور ^{۲۱} ، شعاع بزرگ‌ترین بیضی محصور ^{۲۲} و طول مرز ^{۲۳} |

۳-۴- انتخاب داده‌های آموزشی

برای انتخاب داده‌های آموزشی و آزمایشی، منطقه مورد مطالعه دو قسمت شد؛ به طوری که در هر منطقه، تنوعی از تمامی عارضه‌های شهری موجود باشد. از منطقه مطالعاتی اول، که ۶۰٪ کل منطقه را تشکیل می‌دهد، برای آموزش و اعتبارسنجی و از منطقه مطالعاتی دوم، که ۴۰٪ بقیه منطقه را شامل می‌شود، برای ارزیابی صحت نتایج استفاده شد. این کار به این دلیل انجام شد که در فرایند آزمایش، از داده‌هایی غیر از داده‌های

1. Contrast
2. Dissimilarity
3. Variance
4. Entropy
5. Mean
6. Correlation
7. Homogeneity
8. Second Moment
9. Border Index
10. Asymmetry
11. Compactness
12. Density
13. Elliptic Fit
14. Circular
15. Shape Index
16. Rectangular-Fit
17. Main Direction
18. Length/Width
19. Area
20. Width
21. Radius of Smallest Enclosing Ellipse
22. Radius of Largest Enclosing Ellipse
23. Border Length
24. Classification Tree
25. Non Parametric



شکل ۳. مناطق آموزشی و آزمایشی در محدوده مورد مطالعه

۳-۵-۲- الگوریتم جنگل تصادفی

این طبقه بندی کننده جزء طبقه بندی کننده های گروهی^۹ است و به صورت ترکیبی از درختان طبقه بندی کننده متفاوت عمل می کند. هر درخت تصمیم گیری با استفاده از بخشی از کل ویژگی ها، که تصادفی انتخاب شده اند، ایجاد می شود و در مرحله آخر، نتایج تمامی درختان تصمیم گیری ایجاد شده با هم ترکیب می شوند. این طبقه بندی کننده از دو تکنیک قدرتمند یادگیری ماشین، شامل بگینگ^{۱۰} و انتخاب ویژگی های تصادفی، استفاده می کند. در بگینگ، هریک از درخت ها با بخشی از نمونه های داده های آموزشی آموزش می بیند و پیش بینی ها با رأی اکثریت درختان ساخته می شود. در جنگل تصادفی، علاوه بر انتخاب بخشی از نمونه های آموزشی برای هر درخت تصمیم گیری، به جای استفاده از همه ویژگی ها، زیرمجموعه ای از ویژگی ها در هر گره

بنابراین، نیازمند فرض خاصی در مورد توزیع داده های ورودی نیست و تفاوت اصلی آن با روش های طبقه بندی پارامتریک نظارت شده در همین مورد است (Safavian & Landgrebe, 1991). درخت تصمیم گیری به صورت رویکردی در طبقه بندی تعریف می شود که به طور بازگشتی، مجموعه ای از داده ها را براساس مجموعه ای از شاخص های تعریف شده در هر شاخه^۱ یا گره^۲ در درخت مورد نظر، به زیرمجموعه های کوچک تر تقسیم بندی می کند. این درخت از یک گره ریشه^۴ (متشکل از کل داده ها)، یک مجموعه از گره های میانی^۵ (انشعابات یا شاخه ها) و یک مجموعه از گره های پایانی^۶ (برگ ها) تشکیل شده است. هر گره، در درخت تصمیم گیری، صرفاً یک گره والد^۷ و دو یا چند گره فرزند^۸ دارد. بنابراین، یک مجموعه داده، براساس تصمیم تعریف شده از طریق این درخت، با تقسیم شدن پی در پی طبقه بندی می شوند و براساس گره برگ که در آن مشاهده قرار می گیرد، برچسب کلاس به هر مشاهده اختصاص می یابد (Friedl & Brodley, 1997). در هر تقسیم، یکی از ویژگی های ورودی که در آن امکان تقسیم بندی داده ها بهتر از سایر ویژگی هاست، استفاده می شود تا نمونه های ورودی در گره های میانی و پایانی تقسیم و دارای انشعاب شوند.

- | | |
|----------------------|--------------------|
| 1. Branch | 2. Node |
| 3. Subdivisions | 4. Root Node |
| 5. Intermediate Node | 6. Terminal Node |
| 7. Parent Node | 8. Descendant Node |
| 9. Ensemble | 10. Baging |

گاما، با استفاده از روش ارزیابی تقاطعی ده‌قسمتی^۴ به‌دست آمد. در محیط نرم‌افزار متلب، با به‌کارگرفتن داده‌های آموزشی و بهینه‌کردن پارامترهای این الگوریتم‌ها، طبقه‌بندی در چهار مقیاس متفاوت و از طریق سه الگوریتم یادشده انجام پذیرفت.

۳-۶- ارزیابی صحت طبقه‌بندی

برای ارزیابی صحت نتایج الگوریتم‌های طبقه‌بندی، از نمونه‌های آزمایشی^۵ استفاده شد. در طبقه‌بندی شیء‌پایه، روش‌های بسیاری به‌کار می‌رود؛ از جمله ارزیابی صحت براساس مساحت^۶، ارزیابی صحت براساس تعداد^۷، ارزیابی صحت براساس تطبیق قطعه‌ها با هم^۸ و ارزیابی صحت براساس مشابهت^۹ (Cai et al., 2018). در این تحقیق، ارزیابی صحت براساس مساحت قطعه‌ها انجام شد و پارامترهای ارزیابی‌شده شامل کامل‌بودن، صحیح‌بودن و کیفیت برای قطعات به‌دست آمد (Heipke et al., 1997). علاوه‌بر پارامترهای ذکرشده، صحت کلی و ضریب کاپا نیز برای منطقه مورد نظر محاسبه شده است.

۴- نتایج و بحث

نتایج حاصل از انتخاب مقیاس‌های بهینه، با استفاده از نمودار نرخ تغییر واریانس محلی، در شکل ۴ نشان داده شده است. با توجه به این شکل، می‌توان مشاهده کرد که در مقیاس‌های پایین‌تر، با افزایش اندازه قطعات، واریانس محلی افزایش می‌یابد و نرخ تغییر هم از

به‌طور تصادفی و هنگام رشد درخت ایجاد و از میان آنها انتخاب می‌شود. جنگل تصادفی فقط نیاز به تعریف دو پارامتر برای ایجاد مدل دارد: تعداد درخت‌های طبقه‌بندی‌کننده مورد نظر (k)؛ تعداد ویژگی‌ها (m). به‌این‌ترتیب، مقدار نهایی کلاس اختصاص‌یافته به هر نمونه با رأی اکثریت درخت‌ها تعیین می‌شود (Breiman, 1996).

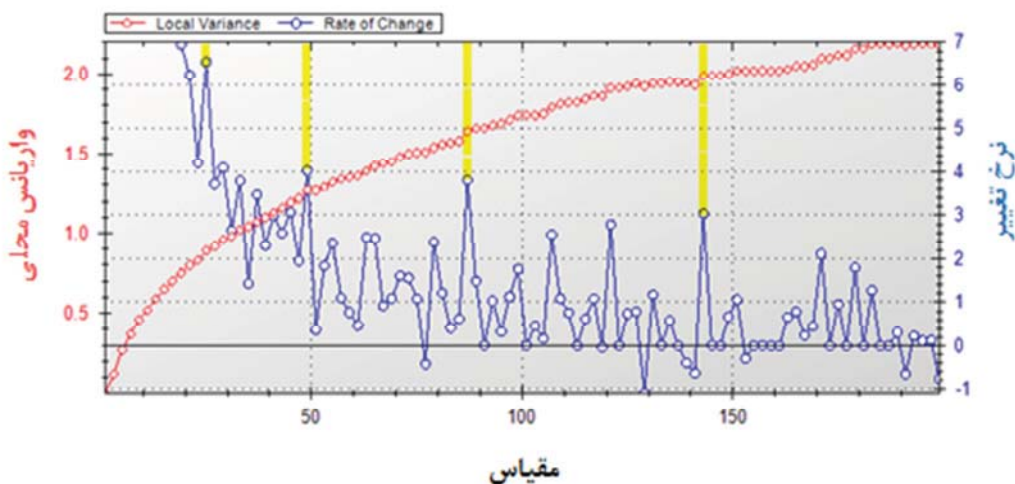
۳-۵-۳- الگوریتم ماشین بردار پشتیبان

ماشین‌های بردار پشتیبان جزء روش‌های یادگیری ماشین به‌شمار می‌آیند و به دو گروه اصلی مدل طبقه‌بندی ماشین بردار پشتیبان و مدل رگرسیون ماشین بردار پشتیبان تقسیم می‌شوند. الگوریتم این طبقه‌بندی را اولین بار ولادیمیر واپنیک^۱، در سال ۱۹۶۳، ابداع کرد و سپس آن را در سال ۱۹۹۵ و به‌همراه کورتس^۲، در مورد حالت‌های غیرخطی تعمیم داد (Cortes & Vapnik, 1995). هدف ماشین‌های برداری پشتیبان طبقه‌بندی داده‌های ورودی به دو کلاس متفاوت است. این روش طبقه‌بندی را به‌صورت باینری انجام می‌دهد؛ (Van der Linden et al., 2009). یکی از محدودیت‌های طبقه‌بندی‌کننده خطی ماشین بردار پشتیبان ناتوانی آن در تفکیک کلاس‌ها در زمانی است که داده‌ها ذاتاً غیرخطی‌اند. در این موارد، رویکردی کلی وجود دارد؛ اینکه داده‌ها را در فضایی با ابعاد بالاتر تصویر کرد تا، در آن فضا، بتوان آنها را از هم تفکیک نمود. در این صورت، امکان کندبودن فرایند یادگیری و کنترل‌ناپذیربودن پیش خواهد آمد. برای رفع این مشکل، از توابع کرنل استفاده می‌شود و بنابراین، با اعمال تابع کرنل، می‌توان نمونه‌هایی را که نیاز به طبقه‌بندی غیرخطی دارند، به‌سادگی، با یک صفحه در فضای جدید از هم تفکیک کرد (Bekkari et al., 2012). به‌منظور اجرای این الگوریتم، از تابع کرنل RBF^۳ استفاده شد و مقدار بهینه، برای پارامترهای تنظیم و

1. Vapnik
2. Cortes
3. Rational Basis Function
4. Cross Validation 10 Fold
5. Test Samples
6. Area-Based Accuracy Measures
7. Number-Based Accuracy Measures
8. Object Matching
9. Feature Similarity-Based Accuracy Measures

نتایج حاصل از انتخاب ویژگی با استفاده از روش جنگل تصادفی در شکل ۵ و جدول ۲ آمده است. در بین ۶۵ ویژگی انتخابی، که شامل چهار گروه ویژگی‌های طیفی و بافتی و ارتفاعی و هندسی بود، تعداد ۲۴ ویژگی به‌منزله مهم‌ترین ویژگی‌ها در جداسازی تاج‌پوشش درختان از سایر عارضه‌های شهری انتخاب شد. طبق شکل ۴، با این تعداد ویژگی، استخراج تاج‌پوشش درختان با بیشترین میزان صحت انجام شده است. میانگین لایه مدل رقومی نرمال‌شده سطح، مدل رقومی سطح، میانگین لایه شیب، میانگین مؤلفه‌های اول و دوم حاصل از آنالیز مؤلفه‌های اصلی و شاخص گیاهی NDVI جزء مهم‌ترین ویژگی‌های انتخابی بودند؛ در حالی که ویژگی‌های هندسی و بافتی در استخراج تاج‌پوشش درختان کمترین تأثیر را داشتند. شباهت درختان با دیگر عارضه‌های شهری، از لحاظ طیفی، جداسازی آنها را با استفاده از اطلاعات طیفی ناممکن می‌کند و استفاده از داده‌های ارتفاعی لیدار در کنار ویژگی‌های طیفی، در جداسازی عارضه‌هایی مانند چمن، که با درختان شباهت طیفی دارند ولی از لحاظ ارتفاعی با آنها متفاوت‌اند، بسیار مؤثر است. در این تحقیق، ویژگی‌های طیفی، در کنار

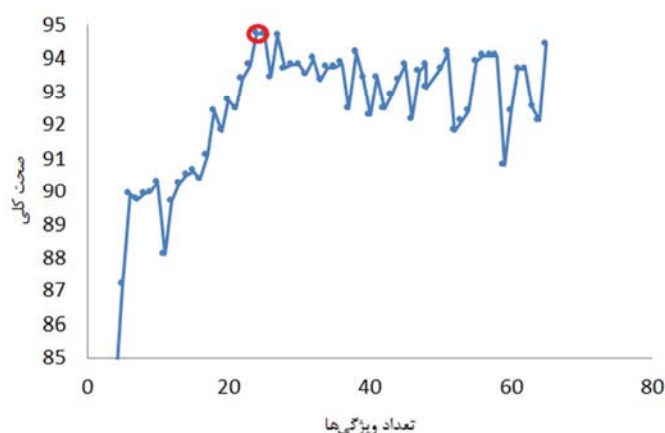
روندی معکوس پیروی می‌کند. این نمودارها به گذار از حالت پیکسل به حالت کوچک‌ترین قطعه معنی‌دار موجود در تصویر اشاره دارد. ملاک تعیین اولین مقیاس اولیه شکستگی در منحنی نرخ تغییر، پس از تنزلی شدید و پیوسته، است. این مقیاس، طبق شکل ۴، مقیاس ۲۵ است. این حد آستانه ممکن است به صورت پله و یا پیک کوچک خود را نشان دهد. با افزایش مقیاس، نوسانات ناگهانی بین پیک‌ها و فرورفتگی‌ها در نمودار نرخ تغییر مشاهده می‌شود؛ در حالی که نمودار واریانس محلی صاف‌تر و نرم‌تر می‌شود. از لحاظ نظری، در نمودار نرخ تغییر، پیک‌ها سطوح مقیاسی را نشان می‌دهند که در آن، به دلیل انطباق قطعات بر عوارض متناظر در دنیای واقعی‌شان، واریانس محلی افزایش می‌یابد. اما باید توجه داشت تغییراتی که با استفاده از قطعه‌بندی پس‌زمینه رخ می‌دهد نیز سبب ایجاد پیک در نمودار نرخ تغییر خواهد شد. با این تفاسیر، در شکل ۴ می‌توان دید که آشکارترین پیک، در قیاس همسایگانش، در مقیاس‌های ۲۵، ۵۰، ۸۷ و ۱۴۳ رخ داده است. به همین علت، در این پژوهش، این چهار مقیاس به‌منزله مقیاس‌های مورد استفاده در فرایند ایجاد قطعات تصویری انتخاب شده‌اند.



شکل ۴. واریانس محلی و نرخ تغییر در مقیاس‌های متفاوت

شیء، با هم ترکیب شده‌اند و یا در مواردی، به صورت خطی قرار گرفته‌اند. این در حالی بود که در موارد دیگر، قطعات فقط شامل تک‌درختان می‌شد و تک‌درختان نیز دارای تنوع شکلی‌اند. بنابراین، نمی‌توان شکل منظم و منحصر به فردی را برای آنها در نظر گرفت تا با سایر عارضه‌های شهری، به صورت کلی، متفاوت باشند. شاید به همین علت ویژگی‌های هندسی، در طبقه‌بندی تاج‌پوشش درختان، از خود عملکرد مناسبی نشان ندادند.

ویژگی‌های ارتفاعی، عملکرد مطلوبی در جداسازی تاج‌پوشش درختان از عارضه‌های دیگر داشته و جنگل تصادفی ویژگی‌های بافتی را از جمله ویژگی‌های مهم در نظر نگرفته است. استفاده از شاخص‌های گیاهی نیز باعث جداسازی عارضه‌هایی مانند ساختمان‌ها، که از لحاظ ارتفاع با درختان شباهت دارند، شده است. تاج‌پوشش درختان دارای شکل‌های متفاوتی است؛ در نتیجه، در برخی مکان‌ها چندین تاج‌پوشش درخت، در قالب یک



شکل ۵. انتخاب ویژگی با روش جنگل تصادفی

جدول ۲. نتایج حاصل از رتبه‌بندی ویژگی‌های انتخابی با استفاده از روش جنگل تصادفی، براساس کاهش میزان خطا

| کاهش میزان خطا | ویژگی | کاهش میزان خطا | ویژگی |
|----------------|--|----------------|---|
| ۲.۴۲ | میانگین لایه مدل رقومی نرمال شده سطح | ۱.۱۰ | انحراف استاندارد مؤلفه اول حاصل از تحلیل مؤلفه‌های اصلی |
| ۱.۸۵ | میانگین لایه مدل رقومی سطح | ۱.۱۰ | انحراف استاندارد مؤلفه دوم حاصل از تحلیل مؤلفه‌های اصلی |
| ۱.۴۴ | میانگین لایه شیب | ۰.۹۶ | انحراف استاندارد مؤلفه سوم حاصل از تحلیل مؤلفه‌های اصلی |
| ۱.۲۸ | میانگین مدل رقومی زمین | ۰.۹۶ | انحراف استاندارد باند قرمز |
| ۱.۲۵ | میانگین مؤلفه اول حاصل از تحلیل مؤلفه‌های اصلی | ۰.۹۰ | انحراف استاندارد باند سبز |
| ۱.۲۵ | میانگین مؤلفه دوم حاصل از تحلیل مؤلفه‌های اصلی | ۰.۹۰ | شاخص پوشش گیاهی NDVI |
| ۱.۲۳ | میانگین باند قرمز | ۰.۸۲ | شاخص پوشش گیاهی DVI |
| ۱.۲۰ | میانگین باند سبز | ۰.۸۰ | شاخص پوشش گیاهی SAVI |
| ۱.۱۸ | انحراف استاندارد لایه شیب | ۰.۷۸ | همگونی باند فرسرخ نزدیک |
| ۱.۱۵ | انحراف استاندارد لایه مدل رقومی سطح | ۰.۷۲ | همبستگی باند قرمز |
| ۱.۱۲ | انحراف استاندارد لایه مدل رقومی نرمال شده سطح | ۰.۶۵ | چگالی |
| ۱.۱۲ | انحراف استاندارد لایه مدل رقومی زمین | ۰.۶ | میانگین مؤلفه دوم حاصل از تحلیل مؤلفه‌های اصلی |

کوچک تر یا درختچه های کم ارتفاع، از این مقیاس نیز می توان در استخراج تاج پوشش درختان استفاده کرد؛ به ویژه، زمانی که باید پردازش های سریعی در زمان های کوتاه انجام گیرد و یا استخراج تاج پوشش ها برای مناطق بسیار گسترده مدنظر باشد. در تحقیقی که مینیت^۱ و همکاران (۲۰۱۱) انجام دادند، کلاس درختان در کوچک ترین مقیاس بهترین شرایط را برای استخراج داشته است و مقیاس های بالاتر برای سایر کلاس ها، مانند ساختمان ها، نتیجه بهتری را نشان داده اند. در تحقیقات دیگری، علت کاهش صحت نتایج طبقه بندی در مقیاس های بالاتر را افزایش *under segment* و *over segment* در جریان قطعه بندی و کاهش تطابق مرزهای قطعات در مقیاس های بالاتر با عارضه های واقعی دانسته اند و در بیشتر موارد، مقیاس های پایین تر نتایج مورد قبول تری داشته است (Liu & Xia, 2010; Zhao et al., 2017). در آگوت^۲ و همکاران (۲۰۱۰)، برای انتخاب مقیاس مناسب در قطعه بندی چند مقیاسه در مورد سه ناحیه مطالعاتی متفاوت، از تغییرات واریانس محلی ناهمگونی قطعات در برابر تغییرات مقیاس استفاده کردند. طی پژوهش مورد نظر، در ناحیه مطالعاتی اول، از داده های مدل رقومی نرمال شده سطح با قدرت تفکیک مکانی ۱ متر، در ناحیه مطالعاتی دوم، از تصاویر با قدرت تفکیک مکانی ۲۵ سانتی متر و در ناحیه مطالعاتی سوم، از تصاویر با قدرت تفکیک مکانی ۶ سانتی متر استفاده شد. قطعات در هر ناحیه، در مقیاس های متفاوتی، با قطعات واقعی به سطح معنی دار رسیدند؛ به صورتی که در ناحیه اول، مقیاس های ۱۴، ۴۵ و ۸۲، در ناحیه دوم، مقیاس های ۱۶، ۳۶ و ۸۶، و در ناحیه سوم مقیاس های ۱۸ و ۳۵ به منزله مقیاس های بهینه انتخاب شدند. نتایج نشان دهنده موفقیت این روش در تعیین مقیاس مناسب، برای قطعه بندی در روش شیء پایه، برای انواع داده ها بود (Drăguț et al., 2010; Myint et al., 2011).

نتایج حاصل از استخراج تاج پوشش درختان در مقیاس های انتخابی با استفاده از سه الگوریتم ماشین بردار پشتیبان، درخت تصمیم گیری و جنگل تصادفی، در جدول ۳ و شکل ۶ نشان داده شده است. صحت کلی، ضریب کاپا، شاخص صحیح بودن، کامل بودن و شاخص کیفیت برای مقیاس های گوناگون محاسبه شده است. شاخص کیفیت برای سطح مقیاس های متفاوت، به ترتیب ۷۹.۱۶، ۷۵.۱۰، ۷۰.۹۳ و ۶۲.۸۶ است. در مقیاس ۲۵ نقشه تاج پوشش درختان با هر سه الگوریتم استفاده شده در این تحقیق، با صحت بالاتری استخراج شده و بنابراین، مناسب ترین مقیاس برای استخراج تاج پوشش درختان مقیاس ۲۵ است. در مقیاس های بالاتر، به دلیل ترکیب تاج پوشش درختان با سایر عارضه های نزدیک از لحاظ طیفی و ارتفاعی، این صحت کاهش می یابد و احتمالاً مقیاس های بالاتر برای استخراج سایر عارضه ها، مانند ساختمان ها، مناسب باشد. با افزایش مقیاس، به دلیل بزرگ شدن قطعات و ترکیب قطعات مربوط به تاج پوشش درختان کوچک با قطعات دیگر، این صحت کاهش می یابد؛ در صورتی که براساس بررسی انجام شده در مورد داده های طبقه بندی شده، تاج پوشش های بزرگ تر در مقیاس های بالا نیز با صحت خوبی استخراج می شوند و این کاهش صحت به دلیل اشتباه در طبقه بندی تاج پوشش های کوچک رخ می دهد.

دلیل استفاده از مقیاس های بزرگ تر افزایش سرعت پردازش و کاهش نویزی شدن نتایج طبقه بندی است. از این رو، در طبقه بندی شیء گرا، تا حد امکان باید از بزرگ ترین مقیاس ممکن برای استخراج عوارض استفاده کرد. با اینکه مقیاس ۲۵، در این تحقیق، بهترین مقیاس در نظر گرفته شده است؛ در مقیاس ۵۰ نیز، نتایج صحت های مورد قبولی دارد و احتمالاً کاهش صحت، در این مقیاس، به علت ترکیب تاج پوشش های کوچک تر با سایر عارضه های مشابه، از جمله چمن یا درختچه های کم ارتفاع، در مرحله قطعه بندی است. بنابراین، در صورت صرف نظر کردن از تاج پوشش های

1. Myint

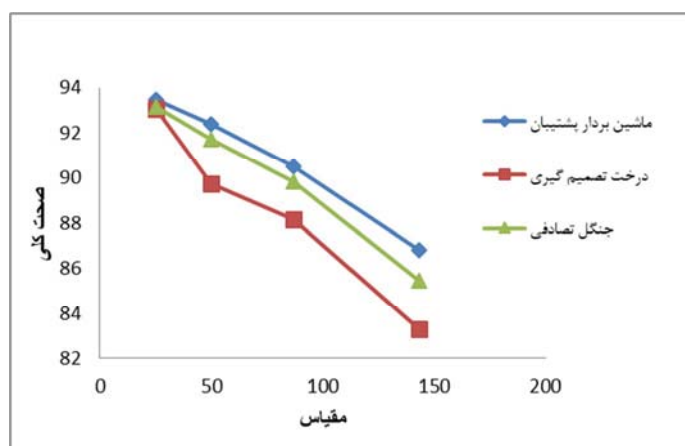
2. Drăguț

همان گونه که در جدول ۳ و شکل ۶ مشاهده می‌شود، الگوریتم ماشین بردار پشتیبان در همه مقیاس‌ها دارای بهترین نتیجه است و الگوریتم درخت تصمیم‌گیری نتیجه‌ای ضعیف‌تر از باقی الگوریتم‌های طبقه‌بندی دارد. الگوریتم جنگل تصادفی نیز، در مقیاس ۵۰ و ۱۴۳، نتایجی ضعیف‌تر از الگوریتم ماشین بردار پشتیبان داشته است. همچنین، نتایج نشان می‌دهد عملکرد این الگوریتم‌ها در کوچک‌ترین مقیاس، یعنی مقیاس ۲۵، به هم نزدیک‌تر است و با افزایش مقیاس، اختلاف بین این الگوریتم‌ها بیشتر می‌شود. به دلیل اینکه پارامترهای هر سه الگوریتم برای

کیم^۱ و همکاران (۲۰۰۸) نیز نشان دادند تعیین پارامتر مقیاس بهینه برای قطعه‌بندی در روش چندمقیاسه، با استفاده از تغییرات واریانس محلی ناهمگونی قطعات، در مورد مناطق جنگلی موفقیت‌آمیز بوده است. در تحقیق دیگری که کاوزوگلو و تانبول^۲ (۲۰۱۸) انجام دادند، قطعه‌بندی چندمقیاسه و انتخاب پارامتر مقیاس با روش انجام‌شده در این تحقیق نتایجی بهتر از دیگر روش‌های قطعه‌بندی داشت. طبق این تحقیقات، می‌توان گفت استفاده از روش تغییرات واریانس محلی ناهمگونی قطعات در برابر مقیاس، که در تحقیق حاضر از آن استفاده شد، نیز نتایج مطمئن‌تری داشته است.

جدول ۳. نتایج طبقه‌بندی با استفاده از الگوریتم‌های یادگیری ماشین، در مقیاس‌های متفاوت

| کیفیت | کامل بودن | صحیح بودن | ضریب کاپا | صحت کلی | مقیاس | طبقه‌بندی‌کننده |
|-------|-----------|-----------|-----------|---------|-------|---------------------|
| ۷۶.۹۰ | ۸۴.۴۱ | ۸۹.۶۲ | ۰.۸۲۲ | ۹۳.۰۷ | ۲۵ | درخت تصمیم‌گیری |
| ۷۱.۱۱ | ۸۲.۸۸ | ۸۳.۴۰ | ۰.۷۵۷ | ۸۹.۷۴ | ۵۰ | |
| ۶۴.۹۹ | ۸۱.۱۸ | ۷۲.۵۵ | ۰.۷۰۶ | ۸۸.۱۶ | ۸۷ | |
| ۵۵.۳۶ | ۷۴.۴۴ | ۶۸.۳۵ | ۰.۵۹۵ | ۸۳.۳۰ | ۱۴۳ | جنگل تصادفی |
| ۷۹.۱۶ | ۹۰.۸۴ | ۸۶.۰۲ | ۰.۸۳۵ | ۹۳.۱۴ | ۲۵ | |
| ۷۵.۱۰ | ۸۹.۳۷ | ۸۲.۴۷ | ۰.۷۹۹ | ۹۱.۷۲ | ۵۰ | |
| ۷۰.۹۳ | ۸۳.۹۲ | ۸۲.۰۸ | ۰.۷۵۲ | ۸۹.۸۱ | ۸۷ | ماشین بردار پشتیبان |
| ۶۲.۸۶ | ۷۳.۴۸ | ۸۱.۳۰ | ۰.۶۶۵ | ۸۵.۴۵ | ۱۴۳ | |
| ۷۹.۹۰ | ۹۲.۱۴ | ۸۵.۷۴ | ۰.۸۴۲ | ۹۳.۴۷ | ۲۵ | |
| ۷۷.۱۴ | ۸۹.۴۵ | ۸۴.۸۶ | ۰.۸۱۷ | ۹۲.۳۸ | ۵۰ | |
| ۷۲.۴۰ | ۸۲.۷۳ | ۸۶.۶۸ | ۰.۷۷۷ | ۹۰.۴۸ | ۸۷ | |
| ۶۴.۶۵ | ۶۴.۶۵ | ۷۹.۹۵ | ۰.۶۹۰ | ۸۶.۴۸ | ۱۴۳ | |



شکل ۶. نتایج طبقه‌بندی با استفاده از الگوریتم‌های یادگیری ماشین در مقیاس‌های متفاوت

شائو و لونتا^۱ (۲۰۱۲) نیز برتری ماشین بردار پشتیبان را در مقایسه با درخت تصمیم گیری، هنگام استفاده از تعداد داده های آموزشی محدود، تأیید می کنند. مقایسه الگوریتم های جنگل تصادفی و ماشین بردار پشتیبان در مطالعات دیگر نیز نشان می دهد که به صورت کلی، ماشین بردار پشتیبان نتایجی بهتر از الگوریتم های دیگر، از جمله جنگل تصادفی و حداکثر احتمال و شبکه عصبی، دربر داشته است (Nitze et al., 2012) که خود این نکته نتایج به دست آمده در تحقیق حاضر را تأیید می کند.

۵- نتیجه گیری

این مطالعه با هدف تهیه نقشه تاج پوشش درختان شهری، با صحت بالا، با ترکیب روش های طبقه بندی شیء پایه و الگوریتم های ماشین بردار پشتیبان، جنگل تصادفی و درخت تصمیم گیری و مقایسه این الگوریتم ها در مقیاس های متفاوت انجام شد. مهم ترین نتایج این تحقیق را می توان این گونه بیان کرد:

- در بین مقیاس های مقایسه شده در این تحقیق، مقیاس ۲۵ دارای بهترین نتیجه در استخراج تاج پوشش درختان است که انتخاب آن تأثیر بسزایی در افزایش صحت نتایج استخراج تاج پوشش درختان داشت؛
- با افزایش مقیاس، صحت استخراج تاج پوشش درختان کاهش می یابد و این به دلیل طبقه بندی نادرست تاج پوشش های کوچک تر و ترکیب آنها با سایر عوارض مشابه، از جمله چمن، است؛
- در جداسازی تاج پوشش درختان از عوارض دیگر، ویژگی های ارتفاعی لیدار و ویژگی های طیفی مهم تشخیص داده شدند؛ به طوری که شاخص گیاهی NDVI، مدل رقومی نرمال شده سطح و میانگین مؤلفه های اول و دوم حاصل از آنالیز مؤلفه های اصلی تأثیرگذارترین ویژگی ها بودند؛
- ویژگی های هندسی قطعات استخراج شده، به دلیل اینکه تاج پوشش گیاهان شکل منظمی نداشتند، در طبقه بندی آنها دارای اهمیتی اندک بود؛

طبقه بندی بهینه شده است و قطعه بندی تصویر، که نتایج طبقه بندی را در روش شیء پایه تحت تأثیر بسیار قرار می دهد، به طور مشابه در مورد هر سه الگوریتم استفاده شده است؛ می توان تفاوت بین الگوریتم ها در مقیاس های بالاتر را ناشی از کاهش تعداد داده های آموزشی دانست که متأثر از افزایش مقیاس است. بدین ترتیب، ماشین بردار پشتیبان و جنگل تصادفی حساسیت کمتری به تغییر در تعداد داده های آموزشی دارند؛ در حالی که درخت تصمیم گیری ناپایداری و حساسیت بالایی در برابر اختلافات در داده های آموزشی دارد و با ایجاد تغییرات کوچک در مقادیر و کیفیت داده های آموزشی، تغییرات چشمگیری در نتایج این الگوریتم حاصل می شود. این ناپایداری و حساسیت بالا از جمله معایب الگوریتم درخت تصمیم گیری به شمار می رود (Timofeev, 2004). اما این مشکل، در جنگل تصادفی، به منزله روش های گروهی با استفاده از چندین یادگیرنده ضعیف و تجمیع نتایج مربوط به این مدل های ایجاد شده، برطرف می شود (Breiman, 2001). این الگوریتم، با استفاده از زیرمجموعه هایی تصادفی از متغیرهای پیشگو، همبستگی بین درختان را کاهش می دهد و سبب بهبود نتایج در مقایسه با زمانی می شود که یادگیرنده های منفرد به کار می روند. بنابراین، گرچه با افزایش مقیاس، صحت نتایج الگوریتم های ماشین بردار پشتیبان و درخت تصمیم گیری کاهش می یابد، این کاهش صحت با شیب ملایمی اتفاق می افتد؛ در حالی که صحت نتایج درخت تصمیم گیری، به دلیل حساسیت بالا به تعداد داده های آموزشی، با افزایش مقیاس کاهش ناگهانی می یابد. این در حالی است که در مقیاس ۲۵، به دلیل وجود داده های آموزشی کافی و حداقل سازی نویز از طریق ایجاد قطعات همگن تر، نتایج سه الگوریتم به یکدیگر نزدیک تر است. تحقیقات گذشته نیز قابلیت بالای الگوریتم ماشین بردار پشتیبان را در طبقه بندی داده های دارای حجم بالا و نمونه های آموزشی اندک، تأیید می کند (Camps-Valls & Bruzzone, 2005; Mallet et al., 2011; Melgani & Bruzzone, 2004).

1. Shao & Lunetta

- Camps-Valls, G. & Bruzzone, L., 2005, **Kernel-based Methods for Hyperspectral Image Classification**, IEEE Transactions on Geoscience and Remote Sensing, 43(6), PP. 1351-1362.
- Chen, M., Su, W., Li, L., Zhang, C., Yue, A. & Li, H., 2009, **Comparison of Pixel-Based and Object-Oriented Knowledge-Based Classification Methods Using SPOT5 Imagery**, WSEAS Transactions on Information Science and Applications, 3(6), PP. 477-489.
- Cortes, C. & Vapnik, V., 1995, **Support-Vector Networks**, Machine Learning, 20(3), PP. 273-297.
- Drăguț, L., Tiede, D. & Levick, S.R., 2010, **ESP: A Tool to Estimate Scale Parameter for Multiresolution Image Segmentation of Remotely Sensed Data**, International Journal of Geographical Information Science, 24(6), PP. 859-871.
- Friedl, M. A. & Brodley, C.E., 1997, **Decision Tree Classification of Land Cover from Remotely Sensed Data**, Remote Sensing of Environment, 61(3), PP. 399-409.
- Gerke, M. & Xiao, J., 2014, **Fusion of Airborne Laserscanning Point Clouds and Images for Supervised and Unsupervised Scene Classification**, ISPRS Journal of Photogrammetry and Remote Sensing, 87, PP. 78-92.
- Heipke, C., Mayer, H., Wiedemann, C. & Jamet, O., 1997, **Evaluation of Automatic Road Extraction**, International Archives of Photogrammetry and Remote Sensing, 32(3 SECT 4W2), PP. 151-160.
- Kavzoglu, T. & Tonbul, H., 2018, **An Experimental Comparison of Multi-Resolution Segmentation, SLIC and K-Means Clustering for Object-Based Classification of VHR Imagery**, International Journal of Remote Sensing, 39(18), PP. 6020-6036.
- Kim, M., Madden, M. & Warner, T. 2008, **Estimation of Optimal Image Object Size for the Segmentation of Forest Stands with Multispectral IKONOS Imagery**, Object-Based Image Analysis (PP. 291-307), Springer.
- مقیاس مورد استفاده در طبقه‌بندی شیء پایه، به دلیل تأثیر در تعداد قطعات ایجاد شده طی فرایند قطعه‌بندی، ممکن است در عملکرد روش مورد استفاده برای طبقه‌بندی قطعات استخراج شده تأثیر بگذارد؛
- روش‌های متفاوت طبقه‌بندی، در مقیاس‌های کوچک‌تر، نتایجی با صحت مشابه به دست می‌دهند؛ در حالی که با افزایش مقیاس، اختلاف عملکرد بین روش‌های طبقه‌بندی افزایش می‌یابد؛
- الگوریتم ماشین بردار پشتیبان به‌منزله بهترین الگوریتم طبقه‌بندی کننده و الگوریتم درخت تصمیم‌گیری به‌منزله ضعیف‌ترین الگوریتم در تمامی مقیاس‌ها شناسایی شدند.

۶- منابع

- Agarwal, S., Vailshery, L., Jaganmohan, M. & Nagendra, H., 2013, **Mapping Urban Tree Species Using Very High Resolution Satellite Imagery: Comparing Pixel-Based and Object-Based Approaches**, ISPRS International Journal of Geo-Information, 2(1), PP. 220-236.
- Baatz, M., 2000, **Multi Resolution Segmentation: An Optimum Approach for High Quality Multi Scale Image Segmentation**. Paper presented at the Beutrage zum AGIT-Symposium. Salzburg, Heidelberg, 2000.
- Bekkari, A., Idbraim, S., Elhassouny, A., Mammass, D., El Yassa, M. & Ducrot, D., 2012, **Spectral and Spatial Classification of High Resolution Urban Satellites Images Using Haralick Features and SVM with SAM and EMD Distance Metrics**, International Journal of Computer Applications, 46(11), PP. 28-37.
- Breiman, L., 1996, **Bagging Predictors**, Machine learning, 24(2), PP. 123-140.
- Breiman, L., 2001, **Random Forests**, Machine Learning, 45(1), PP. 5-32.
- Cai, L., Shi, W., Miao, Z. & Hao, M., 2018, **Accuracy Assessment Measures for Object Extraction from Remote Sensing Images**, Remote Sensing, 10(2), P. 303.

- Liu, D. & Xia, F., 2010, **Assessing Object-Based Classification: Advantages and Limitations**, Remote Sensing Letters, 1(4), PP. 187-194.
- MacFaden, S.W., O'Neil-Dunne, J.P., Royar, A.R., Lu, J.W. & Rundle, A.G., 2012, **High-Resolution Tree Canopy Mapping for New York City Using LIDAR and Object-Based Image Analysis**, Journal of Applied Remote Sensing, 6(1), P. 063567.
- Mallet, C., Bretar, F., Roux, M., Soergel, U. & Heipke, C., 2011, **Relevance Assessment of Full-Waveform Lidar Data for Urban Area Classification**, ISPRS Journal of Photogrammetry and Remote Sensing, 66(6), PP. S71-S84.
- Mathieu, R. & Aryal, J., 2005, **Object-Oriented Classification and Ikonos Multispectral Imagery for Mapping Vegetation Communities in Urban Areas**, 17th Annual Colloquium of the Spatial Information Research Centre (SIRC 2005: A Spatio-temporal Workshop), Dunedin, New Zealand.
- Melgani, F. & Bruzzone, L., 2004, **Classification of Hyperspectral Remote Sensing Images with Support Vector Machines**, IEEE Transactions on Geoscience and Remote Sensing, 42(8), PP. 1778-1790.
- Moussa, A. & El-Sheimy, N., 2012, **A New Object Based Method for Automated Extraction of Urban Objects from Airborne Sensors Data**, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 39, P. B3.
- Myint, S.W., Gober, P., Brazel, A., Grossman-Clarke, S. & Weng, Q., 2011, **Per-Pixel vs. Object-Based Classification of Urban Land Cover Extraction Using High Spatial Resolution Imagery**, Remote Sensing of Environment, 115(5), PP. 1145-1161.
- Niemeyer, J., Rottensteiner, F. & Soergel, U., 2013, **Classification of Urban LiDAR Data Using Conditional Random Field and Random Forests**, Paper presented at the Joint Urban Remote Sensing Event 2013.
- Nitze, I., Schulthess, U. & Asche, H., 2012, **Comparison of Machine Learning Algorithms Random Forest, Artificial Neural Network and Support Vector Machine to Maximum Likelihood for Supervised Crop Type Classification**, Proc. of the 4th GEOBIA, P. 35.
- Parmehr, E.G., Amati, M., Taylor, E.J. & Livesley, S.J., 2016, **Estimation of Urban Tree Canopy Cover Using Random Point Sampling and Remote Sensing Methods**, Urban Forestry & Urban Greening, 20, PP. 160-171.
- Rahman, M. R., & Saha, S. K. (2008). **Multi-resolution segmentation for object-based classification and accuracy assessment of land use/land cover classification using remotely sensed data**. Journal of the Indian Society of Remote Sensing, 36(2), 189-201.
- Safavian, S.R. & Landgrebe, D., 1991, **A Survey of Decision Tree Classifier Methodology**, IEEE Transactions on Systems, Man, and Cybernetics, 21(3), PP. 660-674.
- Shaban, M. & Dikshit, O., 2001, **Improvement of Classification in Urban Areas by the Use of Textural Features: The Case Study of Lucknow City, Uttar Pradesh**, International Journal of Remote Sensing, 22(4), PP. 565-593.
- Shao, Y. & Lunetta, R.S., 2012, **Comparison of Support Vector Machine, Neural Network, and CART Algorithms for the Land-Cover Classification Using Limited Training Data Points**, ISPRS Journal of Photogrammetry and Remote Sensing, 70, PP. 78-87.
- Timofeev, R., 2004, **Classification and Regression Trees (CART) Theory and Applications**, Humboldt University, Berlin.
- Van der Linden, S., Rabe, A., Okujeni, A. & Hostert, P., 2009, **Image SVM Classification**, Application Manual: Image SVM Version, 2.
- Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J. & Atkinson, P.M., 2018, **An Object-Based Convolutional Neural**

Network (OCNN) for Urban Land Use Classification, Remote sensing of environment, 216, PP. 57-70.

Zhao, W., Du, S. & Emery, W.J., 2017, **Object-Based Convolutional Neural Network for High-Resolution Imagery Classification**, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10(7), PP. 3386-3396.



نشن از دور
GIS ایران



سنجش از دور و GIS ایران
Iranian Remote Sensing & GIS

سال سیزدهم، شماره اول، بهار ۱۴۰۰
Vol.13, No. 1, Spring 2021

17-32

Urban Tree Canopy Mapping Using Object Oriented Classification and Machine Learning Algorithms

Haghshenas N.¹, Shamsoddini A.^{2*}, Aghighi H.³

1. M.Sc. Student, Dep. of Remote Sensing and GIS, Tarbiat Modares University, Tehran
2. Assistant Prof., Dep. of Remote Sensing and GIS, Tarbiat Modares University, Tehran
3. Assistant Prof., Dep. of Remote Sensing and GIS, Shahid Beheshti University, Tehran

Abstract

It is necessary to know about the quantity of urban tree canopy cover due to its role in air and noise pollution reduction, wind prevention, saving rain water, and runoff control. Being expensive and time consuming, the manual extraction of tree canopy has been replaced by remote sensing techniques conducted on the images, digitally. There are several parameters which must be optimized prior to use of object oriented classification. One of these parameters is Scale affecting the segmentation results, significantly. Scale is usually set by trial and error which is an experimental approach. One of the aims of this study is to optimize Scale parameter, automatically. In addition, after segmentation process based on a proper Scale, it is required to classify the identified segments based on the attributes which are extracted from these segments. In this stage, the selection of suitable classification method fed by the proper attributes is critical. In this research, LiDAR data and aerial image acquired on Vaihingen, Germany, were utilized for segmenting the urban area. In order to identify suitable attributes, random forest feature selection was applied on the attributes derived from the identified segments. Machine learning methods including support vector machine, random forest, and decision tree were compared for classifying the segments based on their suitable attributes into two classes including tree canopy cover and others. The results indicated that Scale of 25 is the best one to segment this area. Also, the tree canopy cover map derived from support vector machine with quality index of 79.90 showed the best performance among different classifiers used in this study.

Keywords: Urban tree canopy cover, Object oriented method, Scale parameter, Machine learning algorithms.

* Correspondence Address: Tarbiat Modares University, Faculty of Humanities, Dep. of Remote Sensing and GIS, Tehran. Tel: +989155034522

Email: ali.shamsoddini@modares.ac.ir