

ارزیابی قابلیت داده‌های سنجنده GeoEye-1 و پارامترهای بافت تصویر، به منظور

طبقه‌بندی مناظر شهری (مطالعه موردی: منطقه ۳ تهران)

پری گلشنی* - دانشجوی کارشناسی ارشد، گروه جنگلداری، دانشگاه علوم کشاورزی و منابع طبیعی ساری

اصغر فلاح - دانشیار گروه جنگلداری، دانشگاه علوم کشاورزی و منابع طبیعی ساری

جعفر اولادی قادیکلایی - استادیار گروه جنگلداری، دانشگاه علوم کشاورزی و منابع طبیعی ساری

سیاوش کلبی - دانشجوی دکتری، گروه جنگلداری، دانشگاه علوم کشاورزی و منابع طبیعی ساری

پذیرش مقاله: ۱۳۹۱/۱۰/۱۳ تأیید نهایی: ۱۳۹۳/۰۲/۲۷

چکیده

سنجش از دور، برای گردآوری اطلاعات مربوط به تغییرات کاربری در مناطق شهری نقش اساسی دارد. یکی از کامل‌ترین روش‌ها در استفاده از این اطلاعات، طبقه‌بندی است. در میان روش‌های گوناگون طبقه‌بندی، استفاده از آنالیز بافت تصویر برای طبقه‌بندی عوارض انسان‌ساخت، مناسب است؛ زیرا آنالیز بافت نه تنها از اطلاعات طیفی، بلکه از نحوه آرایش فضایی پیکسل‌ها در انجام طبقه‌بندی بهره می‌برد. طبقه‌بندی محیط شهری به دلیل یکنواخت نبودن طبقات، پایین‌ترین صحت را در بین تمام کلاس‌ها دارد، در نتیجه پارامترهای بافت تا حد زیادی می‌تواند صحت تفکیک این کلاس‌ها را افزایش دهد. هدف از این مطالعه، بررسی قابلیت داده‌های سنجنده GeoEye-1، مشخصه‌های بافت تصویر و روش طبقه‌بندی BRT، برای طبقه‌بندی کاربری‌های شهری است. نمونه‌ها از کل منطقه با استفاده از GPS و با توجه به کاربری‌های موجود برداشت شده است که ۳۰ درصد این نقاط برای ارزیابی در نظر گرفته شد. مشخصه‌های بافت تصویر، شامل موارد میانگین، واریانس، آنتروپی، همگنی و عدم تجانس است. در مرحله بعد با استفاده از لایه‌های اصلی و مشخصه‌های بافت تصویر اقدام به طبقه‌بندی شد. نتایج نشان داد این طبقه‌بندی دارای صحت کلی و ضریب کاپا، به ترتیب ۰/۹۲ و ۰/۹۰ است.

کلیدواژه‌ها: آنالیز بافت، الگوریتم طبقه‌بندی BRT، تصویر ماهواره‌ای GeoEye-1، طبقه‌بندی.

مقدمه

استفاده از سنجش از دور در زمینه‌های مربوط به نقشه‌برداری، حفاظت و مدیریت تغییرات زیست‌محیطی و پوشش گیاهی، در توسعه کشورها مقرون به صرفه است. سنجش از دور، برای گردآوری اطلاعات مربوط به پوشش جنگل، انواع

پوشش گیاهی و تغییرات کاربری اراضی، نقش اساسی و محوری ایفا می‌کند. یکی از کامل‌ترین روش‌ها در استفاده از اطلاعات سنجش از دور، طبقه‌بندی است (عزیزی، نجفی، سهرابی، ۲۰۰۸: ۱۱۲۷). طبقه‌بندی، بر اساس تجزیه و تحلیل کیفی اطلاعات برگرفته از مناطق نمونه‌برداری شده، انجام می‌شود. معمولاً روش‌های طبقه‌بندی هنگام ترکیب با ویژگی‌های آماری آنالیز بافت، نتایج مناسبی را ارائه می‌دهد (زنگ، ۱۹۹۹: ۵۰). در واقع می‌توان گفت هنگام استفاده از مشخصه‌های بافت، علاوه بر الگوی طیفی پیکسل‌ها، از الگوی مکانی آنها نیز در طبقه‌بندی بهره می‌بریم (لیلیسند و کيفر، ۲۰۰۴ و لیو، کلارک و هلورد، ۲۰۰۶). برای مثال، نتایج طبقه‌بندی محیط شهری مینت و لام (۲۰۰۵) که با استفاده از تصاویر آیکونوس در شهر نورمن انجام گرفت، نشان داد که طبقه‌بندی با روش‌های معمول طیفی، دارای صحت ۵۵ درصد است و در صورت استفاده از مقادیر مربوط به بافت تصویر این دقت تا ۹۲ درصد افزایش می‌یابد.

در میان خصوصیات مختلف عوارض موجود در یک تصویر (مانند رنگ یا شکل، بافت و...)، بافت بارزترین این عوارض هم از نظر بصری و هم از دیدگاه پردازش تصویر است. بافت یک تصویر به توزیع مکانی مقادیر درجه خاکستری بستگی دارد و شامل اطلاعاتی راجع به کتراست، چین‌خوردگی، نظم و مانند این‌ها است. آنالیز بافت، امکانات جالبی برای تشخیص ناهمگونی ساختاری کلاس‌ها ارائه می‌دهد. بافت ناهمگنی مقادیر مربوط به ارزش پیکسل‌ها در داخل یک منطقه تعریف شده را در تصویر محاسبه می‌کند (وود، پیجن، رادلوف و کیولر، ۲۰۱۲). بافت تصویر برای مشخص کردن الگوهای پوشش گیاهی (جی، کاروتر، جونگ و هررا، ۲۰۰۶) و به‌منزله داده ورودی برای طبقه‌بندی پوشش گیاهی به کار می‌رود. برای مثال می‌توان به طبقه‌بندی کوه‌های راکی کانادا (زنگ و فرانکلین، ۲۰۰۲)، جنگل‌های ساحلی کاناد (کوبرن و روبرت، ۲۰۰۴)، علفزارها و ساوان‌های آفریقا (هوداک و وسمن، ۱۹۹۸) و زیستگاه‌های کوهستانی در آفریقا (است، اکین، وانگی و شوگرت، ۲۰۰۸) اشاره کرد.

علاوه بر این، نتایج مطالعه‌ای که در شهری واقع در اطراف دریاچه تای در شرق چین (یانگ، جیانگ، لو و زنگ، ۲۰۱۲) انجام گرفت، حاکی از آن است که آنالیز بافت در مقایسه با دیگر فرایندهای پردازش، می‌تواند مشخصه‌های عوارض انسان‌ساخت را به‌شکلی معقولانه از تصاویر استخراج کند و به‌منزله مبنایی برای نقشه‌برداری اولیه از توزیع جمعیت روستایی به کار رود.

کایتاکایر، حامل و دیفورنی (۲۰۰۳) بازیابی متغیرهای ساختاری جنگل را با استفاده از آنالیز بافت تصویر سنجنده IKONOS انجام دادند. آنها مشخصه‌های بافت مشتق شده از روش ماتریس سطح خاکستری هم‌اتفاقی (GLCM)^۱ را برآورد کردند. ضریب تبیین (R^2) برای ارتفاع، محیط دور، تراکم توده و سن، دامنه‌ای از ۰/۷۶ تا ۰/۸۲ بود، اما سطح مقطع با مشخصه بافت رابطه ضعیفی داشت ($R^2 = ۰/۳۵$). مشخصه بافت که از ماتریس سطح خاکستری هم‌اتفاقی (GLCM) به‌دست آمده، بیشتر برای طبقه‌بندی و قطعه‌بندی بافت استفاده می‌شود.

متیو و همکارانش با استفاده از تصویر چند طیفی آیکونوس، مطالعه‌ای را برای طبقه‌بندی باغ‌های خصوصی شهری دیوندین^۲ در نیوزلند انجام دادند. در این مطالعه از رویکرد طبقه‌بندی موضوعی متکی بر اطلاعات چندطیفی و اطلاعات

1. Gray Level Co – occurrence Matrix

2. Dunedin

فضایی موجود در تصاویر دیجیتال، استفاده شده است. میانگین صحت طبقه‌بندی برای زمانی که تصویر به شش طبقه تقسیم شد، رضایت‌بخش و ۷۷/۵ درصد محاسبه شد. همچنین ضریب کاپا برای این طبقه‌بندی ۰/۶۹ به‌دست آمد. زمانی که تعداد طبقات به ۴ طبقه کاهش پیدا کرد، متوسط صحت برای طبقات ۹۰/۷ و ضریب کاپا ۰/۸۵ به‌دست آمد (متیو، فریمن، آریال، ۲۰۰۷).

تاپا و مورایاما (۲۰۰۹) تصاویر ALOS^۱ مربوط به شهر تسوکوبای ژاپن را با استفاده از روش‌های نظارت‌نشده (ISODATA)^۲، طبقه‌بندی نظارت‌شده (حداکثر احتمال)، طبقه‌بندی نظارت‌شده فازی و روش پس‌پردازش GIS^۳، به هفت دسته طبقه‌بندی کردند. در این مطالعه روش پس‌پردازش GIS دارای بالاترین صحت کلی (۸۹/۳۳ درصد) بود.

در مطالعه‌ای دیگری که رادریگز و همکارانش در اسپانیا انجام دادند، الگوریتم رندوم فارست^۴ برای طبقه‌بندی کاربری و پوشش زمین مورد آزمون و بررسی قرار گرفت. ارزیابی بر اساس معیارهای صحت نقشه‌برداری، حساسیت به اندازه مجموعه داده‌ها و خطاها بوده است. در این مطالعه از تصاویر ماهواره‌ای بهار و تابستان لندست ۵ تی. ام. برای طبقه‌بندی چهارده نوع کاربری در جنوب اسپانیا استفاده شد. نتایج نشان داد که الگوریتم طبقه‌بندی‌کننده رندوم فارست، طبقه‌بندی را با صحت کلی ۹۲ درصد و با ضریب کاپای ۰/۹۲ محاسبه می‌کند (رادریگز، گیمر و روگان، ۲۰۱۲).

در اکثر مطالعات بیان‌شده، از روش‌های خطی برای طبقه‌بندی داده‌ها استفاده شده است. این روش‌ها دارای یک سری پیش‌فرض از جمله نرمال بودن داده‌ها هستند که در صورت عدم نرمال بودن داده‌ها، استفاده از این روش‌ها برای طبقه‌بندی، خالی از اشکال نیست. در این مطالعه از الگوریتم طبقه‌بندی و رگرسیون BRT^۵، برای طبقه‌بندی کاربری‌های شهری استفاده شده است. این روش از چندین مزیت برخوردار است، از جمله اینکه در مقایسه با روش‌های دیگر طبقه‌بندی، حساسیت کمی به همپوشانی دارد (وستریچ، لسر و فانک، ۲۰۱۰). این روش می‌تواند داده‌های با حجم زیاد را به سرعت اجرا کند، ضمن آنکه نیاز به فرض‌های توزیع داده‌ها ندارد.

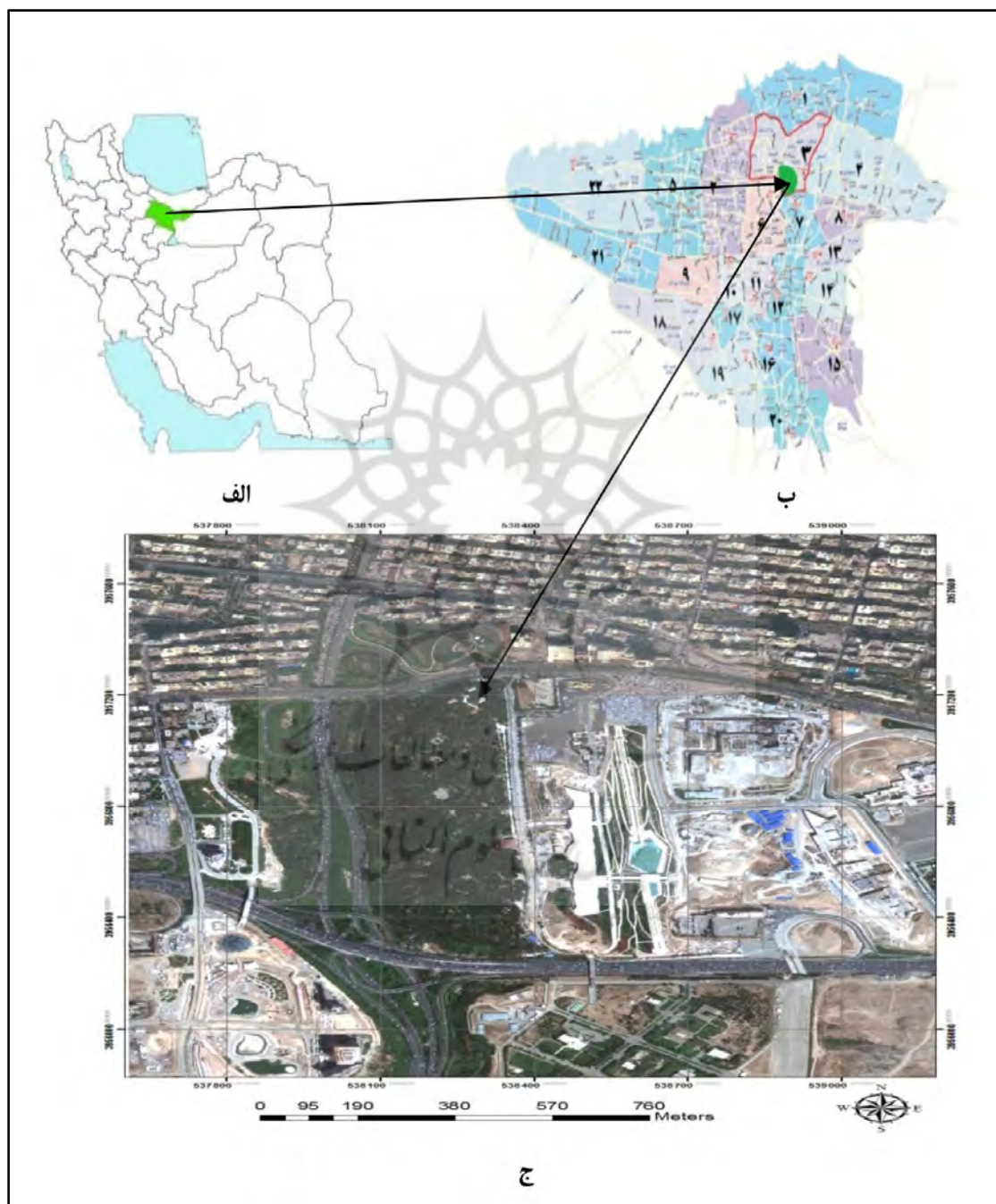
بنابراین استفاده از بافت تصویر به یک جنبه مهم در آنالیز اطلاعات سنجش از دور با وضوح بالا تبدیل شده است (کوبرن و روبرت، ۲۰۰۴ و فرانکلین، وولدر و گریلو، ۲۰۰۱). هدف از این مطالعه، بررسی قابلیت داده‌های سنجنده GeoEye، مشخصه‌های بافت و روش طبقه‌بندی و رگرسیون BRT به‌منظور طبقه‌بندی کاربری‌های شهری است. این اطلاعات در صورت دقت بالا، می‌توانند کارایی بسیار مفیدی برای مدیریت و برنامه‌ریزی مناطق شهری داشته باشند.

منطقه مورد مطالعه

منطقه‌ای که در این پژوهش به بررسی آن پرداخته می‌شود، بخشی از منطقه ۳ شهر تهران است (شکل ۱). پوشش درختی، چمن، جاده، ساختمان‌ها و فضاهای خالی برخی از کاربری‌های مربوط به این منطقه هستند. از جمله کاربری‌های

1. Advanced Land Observing Satellite
2. Iterative Self-Organizing Data Analysis Technique
3. GIS post-processing approach
4. Random Forest
5. Boosted Regression Tree

غالب این منطقه، می‌توان به پارک جنگلی طالقانی اشاره کرد. مختصات جغرافیایی منطقه در سیستم UTM^۱ بدین شرح است: ۵۳ درجه و ۹۱ دقیقه و ۳۲ ثانیه تا ۵۳ درجه و ۷۶ دقیقه و ۸ ثانیه طول شرقی و ۳۹ درجه و ۵۷ دقیقه و ۷۱ ثانیه تا ۳۹ درجه و ۵۶ دقیقه و ۲ ثانیه عرض جغرافیایی.



شکل ۱. تصویر (GeoEye-1) و موقعیت جغرافیایی منطقه مورد مطالعه در ایران؛ الف و ب: استان تهران؛ ج: منطقه ۳

1. Universal Transverse Mercator

مواد و روش‌ها

داده‌های سنجنش از دور و مقادیر مربوط به مشخصه‌های بافت

به‌منظور تهیه نقشه کاربری منطقه، تصاویر ماهواره‌ای GeoEye-1 منطقه مورد مطالعه از سازمان فناوری اطلاعات و ارتباطات شهرداری تهران تهیه شد. این تصاویر در تیر ماه سال ۱۳۸۹ از منطقه گرفته شده است. GeoEye-1 اولین خروجی از نسل جدید سنسورهای ماهواره‌ای تجاری با قدرت تفکیک مکانی بالا است که ششم سپتامبر سال ۲۰۰۸ از کالیفرنیا به فضا پرتاب شد (دنیسن، برانل و کارتر، ۲۰۱۰). این سیستم دارای یک سنجنده پانکروماتیک با قدرت تفکیک مکانی ۴۱ سانتی‌متر و یک سنجنده چند طیفی با قدرت تفکیک ۱/۶۵ متر است. سنجنده طیفی دارای باندهای آبی، سبز، قرمز و مادون قرمز نزدیک است (جدول ۱). مطابق قوانین دولت آمریکا، این تصاویر با پیکسل ۰/۵ متر به مشتریان عرضه می‌شود.^۱ سازمان مربوطه عملیات زمین‌مرجع و تصحیح‌های لازم روی تصاویر را اعمال کرده بود که برای اطمینان از صحت تصحیح هندسی تصویر، از نقشه وکتوری جاده‌ها با مقیاس ۱:۲۰۰۰ استفاده شد. انطباق دقیق لایه جاده‌ها با تصویر گویای صحت تطابق هندسی تصویر است.

جدول ۱. خصوصیات ماهواره GeoEye-1

ارتفاع مدار ماهواره	KM ۶۸۱
عرض نوار برداشت	KM ۱۵ / ۲
طول دوره مأموریت	پیش‌بینی: بیش از ۱۰ سال
زمان عبور از استوا	۱۰:۳۰
تعداد باندها	۵
قدرت تفکیک مکانی برای باندهای رنگی	m ۱/۶۵
رزولوشن برای باند پانکروماتیک	cm ۴۱
محدوده طیفی برای باند پانکروماتیک	nm ۴۵۰ - ۹۰۰
	آبی: ۴۵۰ - ۵۲۰ nm
	سبز: ۵۲۰ - ۶۰۰ nm
محدوده طیفی برای باندهای رنگی	قرمز: ۶۲۵ - ۶۹۵ nm
	مادون قرمز نزدیک ۷۶۰ - ۹۰۰ nm

انواع مختلف روش‌های استخراج بافت را می‌توان به چهار دسته روش‌های آماری، روش‌های ساختاری، روش‌های مبتنی بر مدل و روش‌های تبدیل تقسیم کرد. ماتریس هم‌وقوع درجات خاکستری (GLCM) و ماتریس وقوع درجات خاکستری (GLOM) که در این مطالعه استفاده شده است، یکی از روش‌های متداول استخراج پارامترهای بافت از تصاویر است. در ENVI پنج نوع فیلتر وقوع^۲ (Data range, Mean, Variance, Entropy, Skewness) و هشت نوع فیلتر هم‌وقوع^۳ (Mean, Variance, Entropy, Homogeneity, Dissimilarity, Contrast, Second)

1. <http://Landinfo.com/geo.htm>

2. Occurrence-based filters

3. Co-occurrence-based filters

(moment, Correlation) را می‌توان محاسبه کرد. مطابق نتایج یانگ و همکاران، استفاده از فیلتر وقوع واریانس با ابعاد پنجره ۳×۳، می‌تواند در آشکارسازی عوارض مصنوعی مانند ساختمان‌ها و جاده‌ها مفید باشد (یانگ، جیانگ، لو و زنگ، ۲۰۱۲). در این مطالعه مقادیر واریانس و میانگین با استفاده از اولین مرتبه آنالیز بافت و مقادیر آنتروپی، همگنی و عدم تجانس با استفاده از دومین مرتبه آنالیز بافت محاسبه شده‌اند (جدول ۲). این اطلاعات بیشتر به نوع تصویر آنالیز شده با توجه به دامنه طیفی، قدرت تفکیک سنجنده و خصوصیات شیء مورد مطالعه (دیمانسیون، شکل و توزیع مکانی) بستگی دارد. بافت تصویر برای هر پیکسل با استفاده از نرم‌افزار ENVI محاسبه شد. در پژوهش پیش رو از پنجره ۳×۳ برای همه آنالیزهای بافت استفاده شد؛ زیرا این اندازه از پنجره دارای مزیت استخراج مقادیر پیکسل‌ها در محدوده کوچک است (وود، پیجن، رادلوف و کیولر، ۲۰۱۲). مقادیر عددی مربوط به پیکسل‌ها برای هر تصویر حاصل از آنالیز بافت و همچنین باندهای اصلی تصویر استخراج شدند و پس از آن، داده‌ها با روش BRT طبقه‌بندی شدند.

جدول ۲. روابط آماری آنالیز بافت و نحوه عملکرد آنها در ارتباط با بافت تصویر

فرمول آماری	توصیف آماری	اولین مرتبه آماری
$\sum_{i,j=0}^{N-1} p(i,j)$	میانگین سطح خاکستری در پنجره GLOM	Mean
$\sum_{i,j}^{N-1} p_{i,j}(i, \mu i)^2$	واریانس سطح خاکستری در پنجره GLOM	Variance
دومین مرتبه آماری		
$\sum_{n=0}^{N-1} n \left\{ \sum_{i=1}^N \sum_{j=1}^N p(i,j) \right\}$	اندازه‌گیری مقدار تغییرات محلی در ارزش پیکسل‌های مجاور بوده و با Homogeneity رابطه عکس دارد.	Dissimilarity
$-\sum_i \sum_j p(i,j) \log(p(i,j))$	تنوع شانون. زمانی که مقادیر ارزش پیکسل‌های GLCM دارای ارزش‌های متفاوتی است، دارای بالاترین مقدار است.	Entropy
$\sum_i \sum_j \frac{1}{1 + (i-j)^2 p(i,j)}$	اندازه‌ای که همگنی ارزش پیکسل‌های یک تصویر را بیان می‌کند.	Homogeneity

منبع: هارالیک، شانموگام، دینستین، ۱۹۷۳

جدول ۳ باندهای به‌کار رفته در طبقه‌بندی را نشان می‌دهد. همان‌طور که در جدول مشاهده می‌شود، در مرحله اول اطلاعات مربوط به باندهای تصویر بدون اعمال فیلتر استخراج شد و در مرحله دوم با استفاده از ماتریس وقوع درجات خاکستری (اولین مرتبه آماری)، مقادیر میانگین و واریانس برای هر کدام از باندها به‌دست آمد و در نهایت با استفاده از ماتریس هم‌وقوع درجات خاکستری (دومین مرتبه آماری) مقادیر آنتروپی، همگنی و عدم تجانس برای هر کدام از باندها محاسبه و استخراج شد. سپس مقادیر محاسبه‌شده با استفاده از روش BRT طبقه‌بندی شدند که نتیجه این عملیات، تهیه نقشه کاربری شهری بوده است.

جدول ۳. تعداد و نوع باندهای مورد استفاده در طبقه‌بندی

روش	آنالیز بافت	روش	آنالیز بافت
First-order	باند ۱ (بدون استفاده از فیلتر)	Second-order	باند ۳ (بدون استفاده از فیلتر)
	میانگین باند ۱		میانگین باند ۳
	واریانس باند ۱		واریانس باند ۳
	آنتروپی باند ۱		آنتروپی باند ۳
Second-order	همگنی باند ۱	First-order	همگنی باند ۳
	عدم تجانس باند ۱		عدم تجانس باند ۳
	باند ۲ (بدون استفاده از فیلتر)		باند ۲ (بدون استفاده از فیلتر)
	میانگین باند ۲		میانگین باند ۲
Second-order	واریانس باند ۲	First-order	واریانس باند ۲
	آنتروپی باند ۲		آنتروپی باند ۲
	همگنی باند ۲		همگنی باند ۲
	عدم تجانس باند ۲		عدم تجانس باند ۲

داده‌های آموزشی

هدف اصلی مرحله تعیین نمونه‌های تعلیمی، جمع‌آوری اطلاعات و طبقه‌بندی یک مجموعه آماری است که الگوی انعکاس طیفی هر یک از انواع پوشش زمین را برای طبقه‌بندی تصویر ترسیم می‌کند. تعداد پیکسل‌های آموزشی باید دست کم ده برابر تعداد متغیرهایی باشد که در طبقه‌بندی مدل‌های پارامتری مورد استفاده قرار می‌گیرد؛ هر چند نتایج مطالعات متعدد نشان داده که الگوریتم‌های ناپارامتری برای دستیابی به نتایج مطلوب‌تر به تعداد بیشتری از اطلاعات آموزشی نیاز دارند (رادریگز، گیمر، روکان، ۲۰۱۲: ۹۳).

در این مطالعه، مناطق نمونه‌برداری شده شامل ۸۷۳ پیکسل بوده که از کل منطقه با استفاده از GPS^۱ و با توجه به کاربری‌های موجود در منطقه در تیرماه سال ۱۳۹۰ برداشت شده است. انتخاب داده‌های آموزشی برای ارزیابی طبقه‌بندی، به‌صورت کاملاً تصادفی از بین پیکسل‌های نمونه‌برداری شده بوده است. مطابق جدول ۴ در هر طبقه، ۷۰ درصد از پیکسل‌ها به‌منظور طبقه‌بندی به کار رفته و ۳۰ درصد باقی‌مانده نیز برای ارزیابی استفاده شده است.

جدول ۴. تعداد پیکسل‌های نمونه و آموزشی برای هر طبقه

نوع کاربری	پوشش درختی	چمن	ساختمان	جاده	فضای خالی	مجموع
تعداد پیکسل‌های نمونه	۱۶۳	۱۳۰	۲۰۰	۲۲۰	۱۶۰	۸۷۳
تعداد پیکسل‌های آموزشی	۴۹	۳۹	۶۰	۶۶	۴۸	۲۶۲

روش طبقه‌بندی و رگرسیونی BRT

برای پیش‌بینی پیامدهای^۱ مورد مطالعه، می‌توان از مدل‌ها و روش‌های گوناگونی استفاده کرد که عموماً آنها را مدل‌های پیش‌بینی می‌نامند. روش‌های به‌کار گرفته شده باید براساس نوع متغیر (متغیر پاسخ کمی یا متغیر طبقه‌بندی شده) انتخاب شوند. وقتی پیامد مورد استفاده کمی باشد، عمدتاً از روش‌های رگرسیونی یا مدل‌های عمومی رگرسیونی^۲ استفاده می‌شود. روش رگرسیونی چندمتغیره و روش تحلیل تمایزی^۳ نیز در این خانواده قرار می‌گیرند. این روش‌ها نیازمند پیش‌فرض رابطه خطی بین متغیرهای پیش‌بینی‌کننده^۴ با پیامدها هستند. چنانچه متغیر، پاسخ متغیر طبقه‌بندی شده باشد، می‌توان به‌طور معمول از روش‌های رگرسیون لجستیک (دوجمله‌ای یا چندجمله‌ای) یا لگاریتم خطی^۵ استفاده کرد. این روش‌ها نیز به پیش‌فرض رابطه خطی بین متغیرها نیازمندند. برای مثال در رگرسیون لجستیک باید رابطه خطی بین متغیرهای مستقل و لجیت متغیر پاسخ وجود داشته باشد.

روش‌های درخت تصمیم‌گیری^۶ که برخی متون آن را قسمت‌بندی مکرر^۷ می‌نامند، اولین بار در پیش‌بینی و شناسایی بازاریابی محصولات مورد استفاده قرار گرفت. طبقه‌بندی و رگرسیون درختی (CART)^۸ یکی از روش‌های طبقه‌بندی است که نخستین بار برایمن و همکارانش (۱۹۸۴) آن را ابداع و توسعه دادند. این روش می‌تواند در پیش‌بینی پیامدهای کمی (درخت رگرسیونی) یا طبقه‌بندی شده (درخت طبقه‌بندی) استفاده شود. روش طبقه‌بندی و رگرسیون درختی از طریق مجموعه‌ای از شرط‌های منطقی^۹ (به جای رابطه خطی)، پیامد مورد مطالعه را پیش‌بینی یا طبقه‌بندی می‌کند.

در این مطالعه از الگوریتم BRT برای طبقه‌بندی استفاده شده است. این الگوریتم نتایج بسیاری از طبقه‌بندی‌کننده‌های ضعیف را برای شکل‌گیری یک طبقه‌بندی‌کننده قوی ترکیب می‌کند. در این روش، درختان تصمیم با به‌کار بردن روش CART برایمن و همکاران (۱۹۸۴) ساخته می‌شوند. این روش در واقع یک روش آماری ناپارامتری بوده که از رگرسیون درختی برای حل مشکلات مربوط به رگرسیون خطی بهره می‌برد. در کل، آنالیز مربوط به CART از سه راه انجام می‌شود. ابتدا حداکثر درختان ساخته می‌شود و با مراحل هرس و انتخاب بهترین درختان، با استفاده از فرایند اعتباریابی متقاطع ادامه می‌یابند. این روش از چندین مزیت برخوردار است، از جمله اینکه در مقایسه با روش‌های دیگر طبقه‌بندی، حساسیت کمتری به همپوشانی دارد (وستریچ، لس‌لر و فانک، ۲۰۱۰: ۸۳۱)؛ ضمن اینکه می‌تواند داده‌های با حجم زیاد را به سرعت اجرا کند، نیاز به فرض‌های توزیع داده‌ها ندارد و قادر به تعیین مهم‌ترین مشخصه‌ها در طبقه‌بندی است. مزایای گفته‌شده، این روش را برای طبقه‌بندی داده‌های سنجش از دور مناسب می‌کند (جو، چیهاتا،

1. Outcome
2. General regression models
3. Discriminant function analysis
4. Predictors
5. Log – liner model
6. Decision tree methods
7. Recursive partitioning
8. Classification & Regression Tree
9. Logical if – then conditions

مالت، بوکایر و ۲۰۱۱). عملکرد بهتر این الگوریتم، به تنظیم گزینه‌های مربوط به درختان تقویت‌شونده و پارامترهای توقف شاخه‌بندی (رویش) درختان وابسته است. برای گزینه‌های مربوط به درختان تقویت‌شونده، میزان کاهش^۱ به‌منزله وزن خاص، برای درختان انفرادی و درختان تقویت‌شونده یکی از پارامترهای مهم است. مطالعات تجربی نشان داده‌اند که نرخ کاهش ۰/۱ یا کمتر، معمولاً به سمت مدل‌های مناسب گرایش دارند. علاوه بر این برای مجموعه داده‌های کوچک ($n = 500$) این مقدار می‌تواند ۰/۰۵ تنظیم شود و برای داده‌های بزرگتر ($n = 5000$) می‌توان مقدار ۰/۰۵ را در نظر گرفت (شتایی، ویناکر و بابانجد، ۲۰۱۱). با توجه به بالا بودن تعداد داده‌های مورد بررسی در پژوهش پیش رو، مقدار ۰/۰۵ مد نظر قرار گرفت.

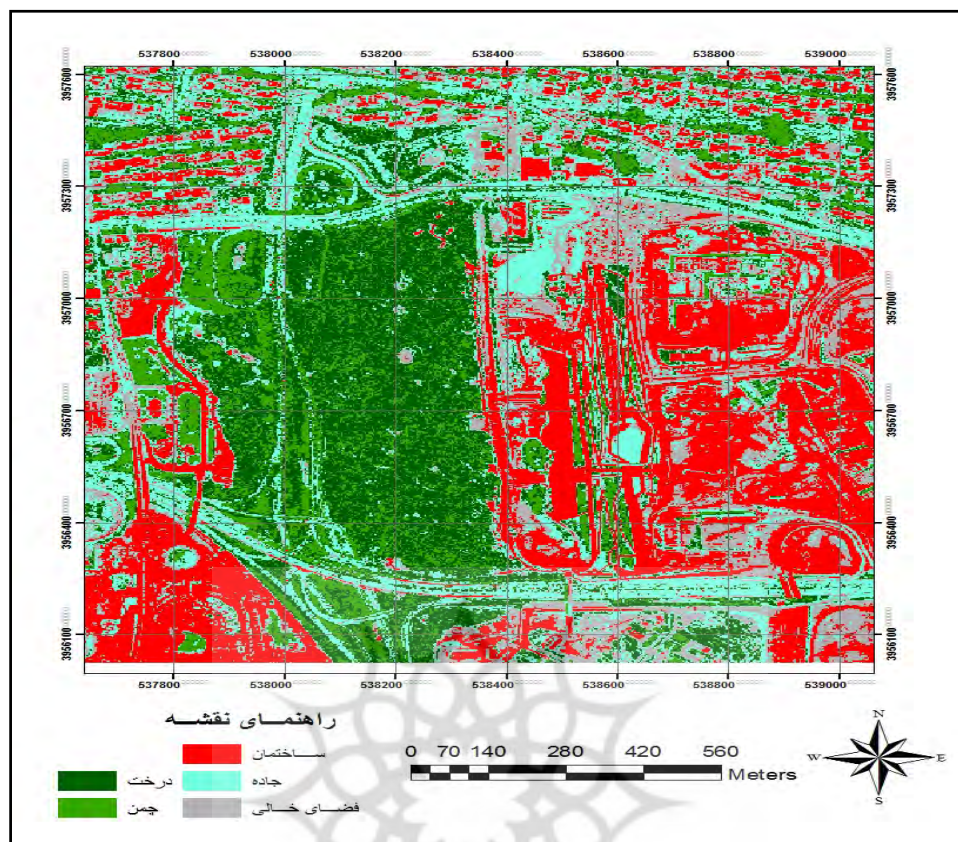
یافته‌های پژوهش

جهت ارزیابی خطاهای مربوط به طبقه‌بندی، معمولاً ماتریس خطای طبقه‌بندی تشکیل می‌شود. گاهی اوقات ماتریس خطا جدول وابستگی یا ماتریس اشتباه نیز نامیده می‌شود. جدول ۵ صحت کلی، صحت کاربر، صحت تولید کننده و ضریب کاپا حاصل از طبقه‌بندی را نشان می‌دهد.

جدول ۵. نتایج جدول خطای طبقه بندی تصویر با استفاده از الگوریتم BRT

نوع کاربری	پوشش درختی	چمن	ساختمان	جاده	فضای خالی	مجموع ردیف
پوشش درختی	۴۰	۸	۰	۱	۰	۴۹
چمن	۳	۳۶	۰	۰	۰	۳۹
ساختمان	۰	۰	۶۰	۰	۰	۶۰
جاده	۲	۰	۴	۶۰	۰	۶۶
فضای خالی	۰	۰	۰	۱	۴۷	۴۸
مجموع ستون	۴۵	۴۴	۶۴	۶۲	۴۷	۲۶۲
صحت تولید کننده	۰/۸۸	۰/۸۱	۰/۹۳	۰/۹۶	۱	-
صحت کاربر	۰/۸۱	۰/۹۲	۱	۰/۹۰	۰/۹۷	-
صحت کلی	۰/۹۲					
ضریب کاپا	۰/۹۰					

نقشه کاربری حاصل از طبقه‌بندی در شکل ۲ قابل مشاهده است.



شکل ۲. نقشه حاصل از طبقه‌بندی تصویر GeoEye-1 مربوط به ناحیه ۴ منطقه ۳ تهران، با استفاده از روش BRT

بحث و نتیجه‌گیری

هدف کلی از طبقه‌بندی تصویر، طبقه‌بندی کلیه پیکسل‌های موجود در تصویر بر حسب نوع پوشش زمینی یا موارد مشابه آن است. برای این کار معمولاً از داده‌های چندطیفی استفاده شده و در حقیقت الگوی طیفی داده‌های هر پیکسل به‌منزله مبنای عددی برای طبقه‌بندی به‌کار گرفته می‌شود. از این رو پدیده‌های مختلف، ترکیبات متفاوتی از ارزش‌های رقومی را نشان می‌دهند و این تفاوت‌ها بر خصوصیات اصلی تابش و انعکاس طیفی آنها مبتنی هستند. با توجه به این مسئله می‌توان گفت که الگوی طیفی یک ویژگی هندسی نیست، بلکه اصطلاح الگو یا طرح، به مجموعه مقادیر پرتوی به‌دست‌آمده از باندهایی با طول موج‌های مختلف برای هر پیکسل اشاره می‌کند. شناسایی الگوی مکانی مستلزم طبقه‌بندی پیکسل‌های تصویر بر مبنای روابط مکانی آنها با پیکسل‌های اطراف است. در طبقه‌بندی مکانی عواملی مانند بافت تصویر، پیکسل مجاور و بافت پدیده‌ها، در نظر گرفته می‌شوند (لیلیسند و کیفیر، ۲۰۰۴). بنابراین می‌توان گفت هنگام استفاده همزمان از باندهای اصلی تصویر و آنالیز بافت در طبقه‌بندی، در واقع از الگوی طیفی و الگوی مکانی پیکسل‌ها به‌صورت همزمان استفاده می‌شود که دارای نتایج بهتری خواهد بود. همان‌طور که قبلاً ذکر شد، آنالیز بافت قادر است صحت طبقه‌بندی را به میزان چشمگیری، به‌خصوص در مناطق ناهمگن شهری ارتقا دهد (مینت و لام، ۲۰۰۵).

همچنین الگوریتم طبقه‌بندی و رگرسیونی BRT یک روش آماری ناپارامتری است که می‌تواند با به‌کارگیری روش CART و بهره‌گیری از درختان تصمیم، طبقه‌بندی را با حجم بالایی از داده‌ها انجام دهد. این روش داده‌های با حجم زیاد را به‌سرعت آنالیز کرده و به فرض‌های توزیع داده‌ها نیاز ندارد، همچنین مهم‌ترین مشخصه‌ها در طبقه‌بندی را تعیین می‌کند. مزایایی که گفته شد، این روش را برای طبقه‌بندی داده‌های سنجش از دور مناسب می‌کند (جو، چیهاتا، مالت و بوکایر، ۲۰۱۱). از این روش در طبقه‌بندی تصاویر ماهواره‌ای استفاده نشده است. با توجه به نتایج به‌دست‌آمده، در این مطالعه استفاده از این روش پیشنهاد می‌شود.

صحت طبقه‌بندی به‌دست‌آمده در این مطالعه ۹۲ درصد محاسبه شد که نسبت به نتایج مطالعه مشابه (متیو، فریمن و آریال، ۲۰۰۷) دارای دقت بالاتری است. از جمله دلایل بالا بودن صحت طبقه‌بندی در این مطالعه، بالا بودن قدرت تفکیک مکانی سنجنده مورد استفاده نسبت به مطالعات مشابه است. مطابق با نتایج هیپا و همکاران، استفاده از سنجنده‌های با قدرت تفکیک مکانی بهتر، می‌تواند میزان خطای برآورد مشخصه‌های کمی جنگل را نسبت به سنجنده‌های دیگر با قدرت تفکیک مکانی کمتر کاهش دهد (هیپا، اینکینن، اینگداهی، لینکو، زو، ۲۰۰۰). از جمله دلایل دیگر بهتر بودن این نتایج نسبت به مطالعات مشابه، استفاده از روش ناپارامتری درخت تقویت‌شونده است و این به‌دلیل قابلیت بالای این روش‌ها در مدل‌سازی و عدم نیاز به نرمال کردن داده‌ها و عدم حساسیت به داده‌های پرت و رابطه همخطی بین متغیرهای مستقل است. در مجموع نتایج این مطالعه نشان داد استفاده از داده‌های سنجنده GeoEye-1 و آنالیز بافت، از قابلیت بالایی در تعیین کاربری اراضی شهری برخوردارند.

سپاس‌گزاری

از سازمان فناوری اطلاعات و ارتباطات شهرداری تهران برای همکاری و در اختیار گذاشتن تصاویر ماهواره‌ای تقدیر و تشکر به عمل می‌آید.

منابع

- Azizi, Z., Najafi, A., Sohrabi, H., 2008, **Forest Canopy Density Estimating, Using Satellite Images, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, Vol. 37, Part B8, PP. 1127-1130.
- Coburn, C. A., Roberts, A. C. B., 2004, **A Multiscale Texture Analysis Procedure for Improved Forest Stand Classification**, International Journal of Remote Sensing, Vol. 25, No. 20, PP. 4287-4308.
- Dennison, P. E., Brunelle, A. R., Carter, V. A., 2010, **Assessing Canopy Mortality during a Mountain Pine Beetle Outbreak Using Geoeye-1-1 High Spatial Resolution Satellite Data**, Remote Sensing of Environment, Vol. 114, No. 14, PP. 2431-2435.
- Estes, L. D., Okin, G. S., Mwangi, A. G., Shugart, H. H., 2008, **Habitat Selection by a Rare Forest Antelope: a Multi-Scale Approach Combining field Data and Imagery from Three Sensors**, Remote Sensing of Environment, Vol. 112, No. 5, PP. 2033-2050.
- Franklin, S. E., Wulder, M. A., Gerylo, G. R., 2001, **Texture Analysis of IKONOS Panchromatic Data for Douglas-fir Forest Age Class Separability in British Columbia**, International Journal of Remote Sensing, Vol. 22, No. 13, PP. 2627-2632.

- Ge, S., Carruthers, R., Gong, P., Herrera, A., 2006, **Texture Analysis for Mapping Tamarix Parviflora Using Aerial Photographs Along the Cache Creek, California**, Environmental monitoring and Assessment, Vol. 114, No. 1-3, PP. 65–83.
- Guo, I., Chehata, N., Mallet, C., Boukir, S., 2011, **Relevance of Airborne Lidar and Multispectral Image Data for Urban Scene Classification Using Random Forests**, ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 66, No. 1, PP. 56–66.
- Haralick, R. M., Shanmugam, K., Dinstein, I. H., 1973, **Textural Features for Image Classification**, IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-3 No. 6, PP. 610–621.
- Hudak, A. T., Wessman, C. A., 1998, **Textural Analysis of Historical Aerial Photography to Characterize Woody Plant Encroachment in South African Savanna**, Remote Sensing of Environment, Vol. 66, PP. 317–330.
- Hyypya, J., H. Hyypya, M. Inkinen, M. Engdahi, S. Linko Y.H. Zhu, 2000, **Accuracy Comparison of Various Remote Sensing Data Sources in the Retrieval of Forest Stand Attributes**, Forest Ecology and Management, Vol. 128, PP. 109-120.
- Kayitakire, F., Hamel, C., Defourny, P., 2000, **Retrieving Forest Structure Variables Based on Image Texture Analysis and IKONOS-2 Imagery**. Remote Sensing of Environment, Vol. 102, PP. 390–401.
- Lillesand, T.M., Kiefer, R.W. & Chipman, J.W., 2004, **Remote Sensing and Image Interpretation**, fifth ed., J. Wiley & Sons, Inc., New York.
- Liu, X. H., Clarke, K., Herold, M., 2006, **Population Density and Image Texture: A Comparison Study**. Photogrammetric Engineering & Remote Sensing, Vol. 72, No. 2, PP. 187-196.
- Mathieu, R., Freeman, C., Aryal, J., 2007, **Mapping Private Gardens in Urban Areas Using Object-Oriented Techniques and Very High-Resolution Satellite Imagery**, Landscape and Urban Planning, Vol. 81, PP. 179–192.
- Myint, S.W., Lam, N., 2005, **A Study of Lacunarity-Based Texture Analysis Approaches to Improve Urban Image Classification**, Computers, Environment and Urban Systems, Vol. 29, PP. 501–523
- Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M., Rigol-Sanchez, J.P., 2012, **An Assessment of the Effectiveness of a Random Forest Classifier for Land-Cover Classification**, ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 67, PP. 93-104.
- Shataee, sh., Weinaker, H., Babanejad, M., 2011, **Plot-level Forest Volume Estimation Using Airborne Laser Scanner and TM Data, Comparison of Boosting and Random Forest Tree Egression Algorithms**, Procedia Environmental Sciences, Vol. 7, PP. 1–11.
- Thapa, R.B., Murayama, Y., 2009, **Urban Mapping, Accuracy, & Image Classification: a Comparison of Multiple Approaches in Tsukuba City, Japan**, Applied Geography, Vol. 29, No. 1, PP. 135–144.
- Westreich, D., Lessler, J., Funk, M. J., 2010, **Propensity Score Estimation: Neural Networks, Support Vector Machines, Decision Trees (CART), and Meta-Classifiers as Alternatives to Logistic Regression**, Journal of Clinical Epidemiology, Vol. 63, PP. 826-833
- Wood, E. M., Pidgeon, A. M., Radeloff, V.C., Keuler, N. S., 2012, **Image Texture as a Remotely Sensed Measure of Vegetation Structure**, Remote Sensing of Environment, Vol. 121, PP. 516–526
- Yang, X., Jiang, G.M., Luo, X., Zheng, Z., 2012, **Preliminary Mapping of High-Resolution Rural Population Distribution Based on Imagery from Google Earth: a Case Study in the Lake Tai Basin, Eastern China**, Applied Geography, Vol. 32, PP. 221-227.
- Zhang, C., Franklin, S. E., 2002, **Forest Structure Classification Using Airborne Multispectral Image Texture and Kriging Analysis**, IEEE, Vol. 4, PP. 2522–2524, DOI: 10.1109/IGARSS.2002.1026598.
- Zhang, Y., 1999, **Optimization of Building Detect on in Satellite Images by Combining Multispectral Classification and Texture Filtering**, ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 54, No. 1, PP. 50-60.