



سجش از دور

GIS ایران

سنجش از دور و GIS ایران / سال هفتم، شماره سوم، پاییز ۱۳۹۴
Iranian Remote Sensing & GIS / Vol.7, No. 3, Autumn 2015

۸۲-۶۵



بررسی بهبود دقت طبقه‌بندی با استفاده از ادغام تصویر تک‌باند ALI با تصاویر ابرطیفی Hyperion

احمد ملک‌نژاد یزدی^{۱*}، حسن قاسمیان^۲، وحید عیسوی^۳، علی شهسواری^۴، حسن کوشا^۵

۱. کارشناس ارشد سنجش از دور و سیستم اطلاعات جغرافیایی، گروه سنجش از دور و GIS، دانشگاه تربیت مدرس
۲. استاد، دانشکده برق و کامپیوتر، دانشگاه تربیت مدرس
۳. کارشناس ارشد سنجش از دور و سیستم اطلاعات جغرافیایی، گروه سنجش از دور و GIS، دانشگاه تربیت مدرس
۴. دانشجوی کارشناسی ارشد سنجش از دور و سیستم اطلاعات جغرافیایی، دانشکده جغرافیا، دانشگاه تهران
۵. کارشناس ارشد سنجش از دور و سیستم اطلاعات جغرافیایی، سازمان نقشه‌برداری ایران

تاریخ پذیرش مقاله: ۱۳۹۴/۱۲/۱۵

تاریخ دریافت مقاله: ۱۳۹۳/۹/۲۸

چکیده

بیشتر الگوریتم‌های طبقه‌بندی داده‌های سنجش از دور براساس ویژگی‌ها و اطلاعات طیفی پیکسل‌ها عمل می‌کنند. این مسئله باعث نادیده گرفتن اطلاعات مکانی سودمند و قابل استخراج بسیاری، مانند بافت تصاویر می‌شود. محیط شهری بافت ناهمگنی دارد که شناسایی انواع کاربری‌ها را به فرایندی دشوار و پیچیده تبدیل کرده است. در این پژوهش تأثیر استفاده از بافت تصویر تک‌باند سنجنده ALI (Advanced Land Imager) بر دقت طبقه‌بندی تصاویر ابرطیفی سنجنده هایپرین «Hyperion» در محیط‌های شهری بررسی شد. طبقه‌بندی با استفاده از روش جنگل‌های تصادفی^۱ و در پنج سناریوی مختلف انجام شد:

سناریوی شماره ۱: طبقه‌بندی اطلاعات طیفی تصویر ادغام‌شده به روش CNT^۱ (بدون لحاظ کردن اطلاعات بافت)؛

سناریوهای ۲، ۳، ۴ و ۵: طبقه‌بندی تصویر CNT با افزوده شدن بردارهای ویژگی بافت حاصل از روش ماتریس هم‌وقوعی در اندازه‌های پنجره ۳، ۵، ۷ و ۹ هستند.

براساس یافته‌های این تحقیق، افزودن بافت به طیف تصویر ادغامی به روش CNT دقت طبقه‌بندی را بهبود چشمگیری داد، به گونه‌ای که دقت کلی بر اثر افزودن بافت با حدود ده درصد افزایش، از ۸۰.۴۷٪ به ۹۰.۷۴٪ رسید. بسیاری از پوشش‌های کاربری مانند جاده، بافت مسکونی، صنایع کوچک و پراکنده و صنایع متمرکز نیز در زمینه دقت تولیدکننده و مصرف‌کننده رشد چشمگیری را شاهد بودند. پارامتر خطای OOB^۱ با ۱۱٪ کاهش، از ۱۹.۸۶ به ۸.۸۷٪ رسید. بردارهای ویژگی مانند میانگین و کنتراست نیز، توانستند از لحاظ میزان اهمیت در رتبه‌های بالا قرار گیرند. همچنین، افزایش اندازه پنجره منجر به بهبود بیشتر دقت طبقه‌بندی شد، به گونه‌ای که اندازه پنجره ۹ بهترین عملکرد را در پی داشت.

کلید واژه‌ها: سنجش از دور، تصاویر ابرطیفی، ادغام تصاویر، جنگل‌های تصادفی، بافت.

* نویسنده عهده‌دار مکاتبات: مشهد، خیابان ابوذر غفاری، ابوذر غفاری ۱۳، پلاک ۱۱۰/۱، کد پستی: ۹۱۷۶۸۷۴۷۸۴، تلفن ثابت: ۰۵۱۳۲۸۴۰۷۱۰۷

تلفن همراه: ۰۹۱۵۳۲۰۵۵۲۴، دورنگار: ۳۶۵۱۶۰۲۱

۱- مقدمه

نقشه‌های کاربری/پوشش اراضی از نقشه‌های مهم و کاربردی در بسیاری از حوزه‌های تحقیقاتی و عملیاتی همچون برنامه‌ریزی شهری، مطالعات مرتبط با منابع آب، کنترل سیلاب و پایش زیستگاه‌های طبیعی به شمار می‌آیند. تغییرپذیری نسبتاً بسیار مؤلفه‌های شهری، هزینه بالای عملیات نقشه‌برداری زمینی و نیاز به نقشه‌های به‌روز این حوزه کاربردی را به یکی از زمینه‌های کاری جالب توجه برای سنجش از دور تبدیل کرده است (Zhang et al., 2012).

بافت شهری با دارا بودن محیطی ناهمگون، تغییرپذیری نسبتاً بسیار مؤلفه‌های آن، تنوع در مصالح سازنده (بتن، آسفالت، فلزات، پوشش گیاهی و خاک) و گستردگی طیفی آن‌ها در تصاویر، محیطی پیچیده برای کارهای مطالعاتی در حوزه سنجش از دور به شمار می‌آید. وجود این بافت ناهمگون و متنوع و گستردگی طیفی، شناسایی انواع کاربری‌های اراضی شهری را به فرایندی پیچیده و دشوار برای سنجنده‌های چندطیفی تبدیل کرده است. (ibid.)

تصاویر ابرطیفی، مانند تصاویر سنجنده‌های پیریون، با دارا بودن ۲۴۰ باند طیفی می‌توانند در رفع این تنوع و پیچیدگی طیفی مواد تا اندازه‌ای کارآ باشند. تصاویر ابرطیفی سنجنده‌های پیریون برای مطالعات مناطق شهری در برخی موارد استفاده شده است (Kasetkasem et al., 2005; Welikanna et al., 2008). مهم‌ترین مانع در استفاده از این سنجنده در محیط شهری، قدرت تفکیک مکانی پایین ۳۰ متر آن است که برای مطالعات شهری مناسب نیست. بهبود دقت طبقه‌بندی داده‌های این سنجنده مستلزم به‌کارگیری روش‌ها و تکنیک‌های خاصی است که از این تکنیک‌ها می‌شود به ادغام تصاویر ابرطیفی با تصاویر تک‌رنگ دارای دقت مکانی بالا اشاره کرد. این روش را محققان گوناگون مطالعه کرده و به کار گرفته‌اند

(Dobhal, 2008; Pande et al., 2009; Licciardi et al.,)

(Pour et al., 2013; 2012).

رویکرد دیگر در بهبود دقت طبقه‌بندی استفاده از اطلاعات مکانی به منظور افزایش دقت طبقه‌بندی حاصل از اطلاعات طیفی است. از جمله این روش‌ها می‌شود به روش میدان تصادفی مارکوف^۱ و نگاشت با قدرت تفکیک مکانی بسیار بالا^۲ اشاره کرد. استفاده از بافت تصاویر و افزودن آن به اطلاعات طیفی نیز یکی از رویکردهای مورد استفاده در این زمینه است (Rodriguez-Galiano et al., 2012; Kasetkasem) (et al., 2005; Welikanna et al., 2008).

پژوهشگران مختلفی از اطلاعات بافت تصویر برای طبقه‌بندی کاربری و پوشش اراضی استفاده کرده‌اند. برای نمونه، در پژوهشی با استفاده از روش میدان‌های تصادفی مارکوف، افزایش محسوسی در دقت طبقه‌بندی پوشش‌های کاربری حاصل از به‌کارگیری تصاویر سنجنده + ETM ماهواره لندست گزارش شد (Kasetkasem et al., 2005).

در پژوهشی دیگر استفاده همزمان از بافت و اطلاعات طیفی تصاویر سنجنده IKONOS در کاری تحقیقاتی به منظور طبقه‌بندی اجتماعات گیاهی جزیره Heard در نزدیکی قطب جنوب مورد استفاده قرار گرفت. سه سناریوی گوناگون طبقه‌بندی چندطیفی باندهای تصویر، طبقه‌بندی ویژگی‌های بافت و طبقه‌بندی ترکیب بافت و باندهای تصویر طراحی شد. نتایج به دست آمده مؤید برتری و افزایش ۶ درصدی دقت خروجی حاصل از ترکیب بافت با اطلاعات طیفی است (Murray et al., 2010). با توجه به آنچه گفته شد، کاربرد و ادغام اطلاعات مکانی تصویر تک‌رنگ سنجنده ALI با تصاویر ابرطیفی سنجنده‌های پیریون در پژوهش‌های پیشین محققان کمتر مورد توجه قرار

1. Markov random fields
2. super resolution mapping

جنوب اسپانیا، اثر افزودن بافت ناشی از ماتریس هم‌وقوعی در افزایش دقت طبقه‌بندی، با دو الگوریتم جنگل‌های تصادفی و بیشترین میزان مشابهت، بررسی شد. الگوریتم جنگل‌های تصادفی، در مقایسه با طبقه‌بندی‌کننده بیشترین میزان مشابهت متکی بر داده‌های طیفی، توانست با بهره‌گیری از قابلیت‌های بافت حدود ۱۱٪ افزایش دقت کلی طبقه‌بندی را داشته باشد. این افزایش دقت در برخی از کلاس‌ها تا ۳۱٪ مشاهده شد (Rodriguez-Galiano et al., 2012).

با توجه به قابلیت‌های ذکر شده، الگوریتم جنگل‌های تصادفی و همچنین، تهیه همزمان تصویر تک‌رنگ سنجنده ALI و تصویر سنجنده هایپریون، استفاده از بافت تصویر تک‌رنگ سنجنده ALI و الگوریتم جنگل‌های تصادفی، برای بهبود دقت تصاویر ابرطیفی در این تحقیق، هدفگذاری شد. افزون بر این، در این مطالعه، اهمیت هر یک از باندهای طیفی مورد استفاده تصویر هایپریون و ویژگی‌های مکانی تصویر ALI در طبقه‌بندی پوشش اراضی معرفی و بحث خواهد شد.

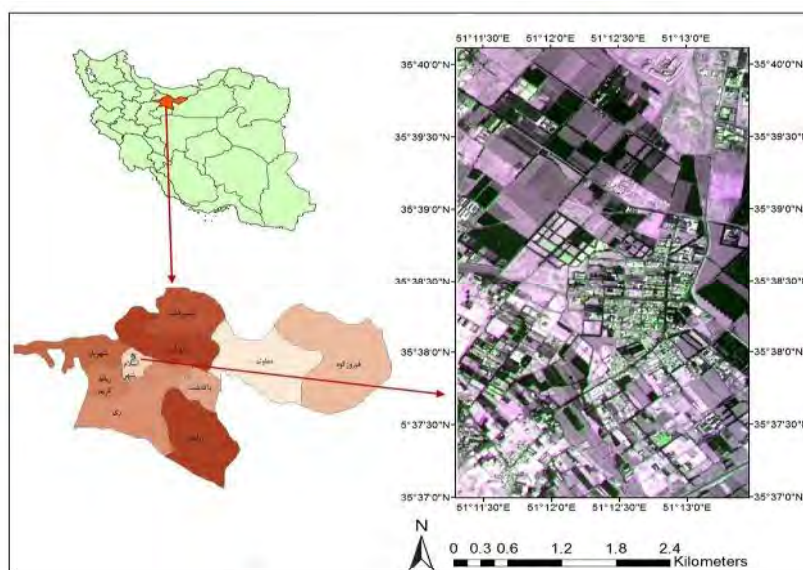
۲- داده و منطقه مورد مطالعه

منطقه مورد مطالعه در این پژوهش بخشی از ناحیه جنوب غرب تهران است که کاربری‌ها و پوشش‌های متنوعی مانند مناطق مسکونی، مناطق صنعتی، باغات، مزارع و زمین‌های بایر دارد. محدوده مورد مطالعه ۱۹۳۸ هکتار از اراضی کشاورزی و سکونتگاهی جنوب غرب تهران، مانند شهرک احمدآباد مستوفی در مجاورت آزادراه‌های آزادگان و خلیج فارس، را در بر می‌گیرد. این منطقه همچنین، در دشتی مسطح و هموار و بدون پیچیدگی‌های توپوگرافی، نظیر شیب‌های تند و مناطق مرتفع، قرار دارد (شکل ۱). تصاویر استفاده شده در این تحقیق، هر دو، از ماهواره EO-1 دریافت شده‌اند. مشخصات سنجنده‌های ALI و هایپریون مستقر بر این سکو در جدول ۱ آمده است.

گرفته است. در صورتی که یکی از قابلیت‌های ماهواره EO-1 که حامل سنجنده‌های ALI و هایپریون است، ارائه تصویر تک‌رنگ با قدرت تفکیک مکانی ۱۰ متر است که به وسیله سنجنده ALI و همزمان با تصویر سنجنده هایپریون دریافت می‌شود.

از مشکلات پیش رو در طبقه‌بندی داده‌های ابرطیفی، بعد زیاد داده‌ها است که به پدیده هیوز می‌انجامد (Chutia et al., 2014; Hsu, 2007). بر اثر این پدیده، تعداد نمونه‌های آموزشی مورد نیاز به صورت توانی با افزایش تعداد باندهای طیفی افزایش می‌یابد. از دیگر سو، تنوع طیفی بالای داده موجب پیدایش نمونه‌های آموزشی با چند مؤلفه طیفی گوناگون می‌شود. این موارد دقت طبقه‌بندی‌کننده‌های پارامتریک را بسیار کاهش می‌دهد (Qiu, 2008) که متکی بر فرض نرمال بودن توزیع نمونه‌های آموزشی است (Lu et al., 2007). همچنین، با توجه به توزیع غیرنرمال ویژگی‌های بافت، ممکن است کاربرد روش‌های غیرپارامتریک مناسب‌تر باشد. الگوریتم جنگل‌های تصادفی با قابلیت‌های خود، از جمله غیرپارامتریک بودن، امکان به کار رفتن در طبقه‌بندی همزمان بافت و داده‌های طیفی را دارد.

الگوریتم جنگل‌های تصادفی گسترش یافته درخت‌های تصمیم‌گیری است (Fawagreh et al., 2001; Breiman, 2014). این الگوریتم با قابلیت مدیریت حجم بالای داده ورودی، توانسته در زمینه بهبود دقت نقشه‌های کاربری و پوشش اراضی عملکرد قابل قبولی از خود ارائه دهد (Ghimire et al., 2010; Rodriguez-Galiano et al., 2012). از ویژگی‌های دیگر این الگوریتم قابلیت برآورد میزان اهمیت متغیرهای مورد استفاده در میزان دقت طبقه‌بندی است که برای سنجش اهمیت نسبی متغیرهای به کار رفته در طبقه‌بندی سودمند واقع می‌شود. در تحقیقی روی پوشش‌های کاربری متنوع مناطق



شکل ۱. موقعیت مکانی منطقه مورد مطالعه

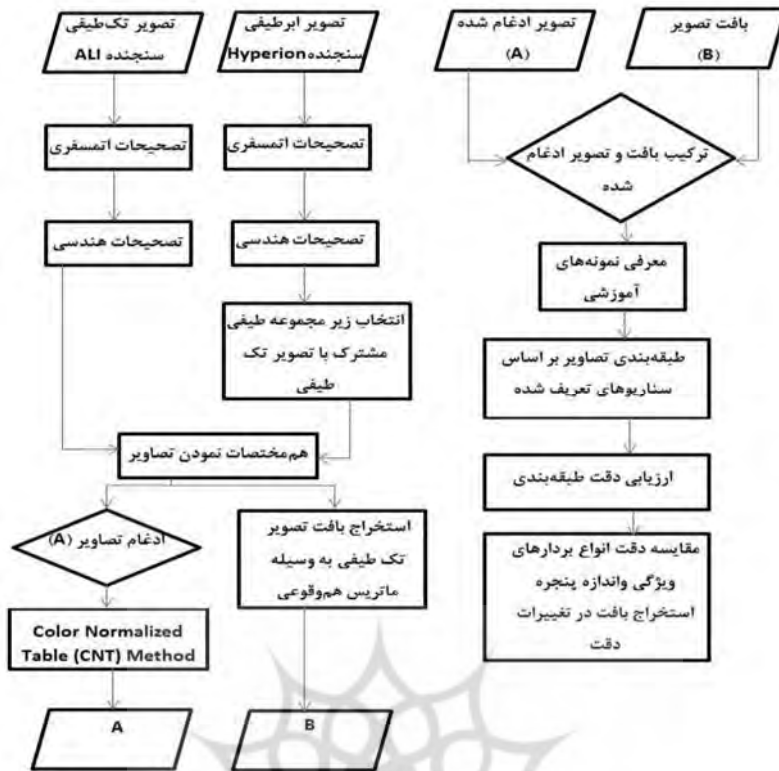
جدول ۱. پارامترهای گوناگون سنجنده‌های ALI و Hyperion

پارامتر مورد بررسی	سنجنده ALI	سنجنده Hyperion
محدوده طیفی	۴۰۰-۲۴۰۰ نانومتر چندطیفی - ۴۸۰ تا ۶۹۰ نانومتر تک‌رنگ	۲۵۰-۴۰۰ نانومتر
قدرت تفکیک مکانی	۳۰ متر	۳۰ متر
قدرت تفکیک طیفی	متغیر	۱۰ نانومتر
عرض تصویربرداری	۳۷ کیلومتر	۷.۶ کیلومتر
قدرت تفکیک باند تک‌رنگ	۱۰ متر	ندارد
تعداد کل باندهای طیفی	۱۰	۲۴۰
تاریخ دریافت تصویر	۸ آبان ۱۳۸۸	۸ آبان ۱۳۸۸

۳- روش تحقیق

مسکونی و صنعتی، باغات، جاده‌ها و معابر و ردیف‌های درخت بودند. پنج سناریوی گوناگون برای طبقه‌بندی تعریف شد: تعریف سناریوی نخست براساس طبقه‌بندی داده‌های صرفاً طیفی حاصل از ادغام تصویر تک‌رنگ و ابرطیفی تعریف بود. در چهار سناریوی باقی‌مانده، اثر افزودن بافت با اندازه‌های پنجره ۳، ۵، ۷ و ۹ در تغییر دقت طبقه‌بندی بررسی شد. با استفاده از روش جنگل‌های تصادفی در محیط نرم‌افزار R طبقه‌بندی صورت گرفت و دقت طبقه‌بندی با استفاده از نمونه‌های آزمایشی ارزیابی شد. در پایان، اثر اندازه پنجره و همچنین، میزان اهمیت بردارهای ویژگی در سناریوهای گوناگون مورد بررسی قرار گرفت.

طرح کلی این تحقیق در شکل ۲ نشان داده شده است. در ابتدای این تحقیق، پردازش‌ها و تصحیحات هندسی و اتمسفری لازم برای آماده‌سازی داده جهت ادغام تصاویر صورت گرفت. سپس بافت تصویر تک‌رنگ سنجنده ALI به وسیله ماتریس هم‌وقوعی استخراج شد. در ادامه، تصویر ابرطیفی با تصویر تک‌رنگ ادغام شد. با استفاده از تصاویر با دقت مکانی بالا و پیمایش‌های میدانی، کلاس‌های موجود در منطقه برای معرفی نمونه‌های آموزشی به طبقه‌بندی‌کننده شناسایی شدند. کلاس‌های عمده موجود در منطقه شامل مزارع ذرت، انواع زمین‌های بایر و کشاورزی، بافت‌های



شکل ۲. نمودار مراحل مختلف تحقیق

۳-۱- پیش پردازش داده

تصاویر ابرطیفی سنجنده EO-1 به علت پهنای باند بسیار کم، بیشتر از تصاویر چندطیفی تحت تأثیر عوامل جوی قرار دارند بنابراین، لازم است با به کارگیری الگوریتم های تصحیح اتمسفری، آثار نامطلوب جوی به کمترین میزان رسانده شود. در مرحله بعد، با حذف باندهای ناسالم تصویر ابرطیفی (باندهای تخریب شده، باندهای جذبی، باندهای کالیبره نشده)، در نهایت ۲۲۰ باند اولیه به ۱۳۴ باند کاهش یافت. سپس الگوریتم FLAASH^۱ برای انجام اصلاحات اتمسفری به کار رفت.

تصحیح هندسی برای هم مختصات کردن بسیار دقیق تصاویر مورد نظر، جهت ادغام، در گام بعد صورت پذیرفت. بدین منظور، با مرجع قرار دادن تصویر تک رنگ، فرایند تثبیت و هم مرجع سازی، با استفاده از ۱۹ نقطه کنترل زمینی در کل تصویر و معادلات چندجمله ای درجه ۲ با دقت ۰/۴۲ پیکسل (۱۲.۶ متر)

انجام شد. برای درون یابی تصویر نیز روش نزدیک ترین همسایه به کار رفت.

۳-۲- استخراج اطلاعات بافت

روش های استخراج بافت تصاویر را می شود به چهار دسته عمده: روش های ساختاری، آماری، استخراج بافت براساس مدل و استخراج بافت براساس تبدیل تقسیم کرد. روش ماتریس هم وقوعی نخستین روش استخراج بافت به روش آماری است که Haralick در سال ۱۹۷۳ آن را ارائه کرد (Haralick et al., 1973). این روش براساس فاصله و زاویه میان دو زوج پیکسل که در یک پنجره با ابعاد مشخص قرار گرفته اند، پارامترهای مرتبط را که بردارهای ویژگی نامیده می شوند، محاسبه می کند. در تحقیق اولیه Haralick، ۲۳ بردار ویژگی

1. fast line of sight atmospheric analysis spectral hypercubes

رابطه (۷)

$$\text{entropy} = - \sum_i^{N-1} \sum_j^{N-1} P_d(i,j) \cdot \log P_d(i,j)$$

رابطه (۸)

$$\text{angular Second Moment} = \sum_i^{N-1} \sum_j^{N-1} P_d(i-j)^2$$

رابطه (۹)

$$\text{correlation} = \frac{\sum_i^{N-1} \sum_j^{N-1} (i,j) P_d(i,j) - \mu_i \mu_j}{\sigma_i \cdot \sigma_j}$$

در این شش رابطه μ و σ ، به ترتیب، میانگین و واریانس محسوب می‌شوند. بردار ویژگی همگنی مقادیر بالاتری را در اختلافات پایین‌تر نشان می‌دهد. این بردار به مؤلفه‌های نزدیک قطر ماتریس هم‌وقوعی حساسیت بیشتری نشان می‌دهد. کنتراست مؤلفه‌ای است که تناوب مکانی را نشان می‌دهد که بیانگر تفاوت میان بیشینه مقادیر و کمینه مجموعه‌ای از پیکسل‌های مجاور است. عدم‌مشابهت، برخلاف کنتراست که به صورت توانی تغییر می‌کند، تغییراتی خطی دارد. این مؤلفه هنگامی که مناطق تصویر کنتراست بالا داشته باشند، مقادیر بالاتری را تجربه می‌کند. آنروپی، به نوعی، بی‌نظمی موجود در تصویر را منعکس می‌کند. هنگام یکنواخت نبودن تصویر، این مؤلفه مقادیر بسیار بالایی دارد. ممان دوم زاویه‌ای معیاری برای یکنواختی بافت تصویر است که، به نوعی، تکرار مقادیر جفتی پیکسل‌ها را نشان می‌دهد. همبستگی نشانگر وابستگی‌های خطی میان مقادیر موجود در تصویر است. مقادیر بالای همبستگی به روابط خطی بین مقادیر سطوح خاکستری در جفت پیکسل‌ها را اشاره دارد (Pacifci et al., 2009).

پارامترهای مؤثر در بافت استخراج شده به وسیله ماتریس هم‌وقوعی، اندازه پنجره به کار رفته در استخراج بردارهای ویژگی و فاصله میان زوج پیکسل‌ها هستند. تغییر این دو پارامتر ممکن است در استخراج

معرفی شد که در سال ۱۹۷۹، او این بردارها را به ۸ بردار ویژگی تقلیل داد (Haralick, 1979). در این پژوهش نیز از این هشت بردار ویژگی استفاده شده است. این هشت بردار، به ترتیب، عبارت‌اند از: میانگین، واریانس، همگنی، کنتراست، عدم‌مشابهت، آنروپی، ممان دوم زاویه‌ای و همبستگی. بردارهای ویژگی میانگین^۱ و واریانس^۲ براساس روابط ۱ تا ۳ تعریف می‌شوند

$$\mu_i = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} i p_d(i,j) \quad (۱)$$

$$\sigma_i^2 = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P_d(i,j) (i - \mu_i)^2 \quad (۲)$$

$$\sigma_j^2 = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P_d(i,j) (j - \mu_j)^2 \quad (۳)$$

در این روابط، i و j تن سطح خاکستری در پنجره محاسبه‌کننده و همچنین، مختصات در فضای ماتریس هم‌وقوعی هستند. $P_d(i,j)$ نیز نشانگر تناوب‌های تکرار نرمال‌شده دو سلول همسایه جداشده با یک فاصله معین است که یکی از آن‌ها تن سطح خاکستری i و دیگری تن سطح خاکستری j دارد. بردار ویژگی میانگین از سطوح خاکستری در پنجره محاسبه‌کننده میانگین‌گیری می‌کند. بردار ویژگی واریانس به محاسبه واریانس سطوح خاکستری در پنجره محاسبه‌کننده می‌پردازد. هنگامی که انحراف معیار استاندارد میان سطوح خاکستری در منطقه‌ای مقادیر بالایی داشته باشند، این پارامتر مقادیر زیادی خواهد داشت. این دو بردار ویژگی دو مؤلفه سطح اول نامیده می‌شوند (Pacifci et al., 2009).

شش بردار ویژگی دیگر که مؤلفه‌های سطح دوم خوانده می‌شوند، همگنی^۳، کنتراست^۴، عدم‌مشابهت^۵، آنروپی^۶، ممان دوم زاویه‌ای^۷، همبستگی^۸ هستند که براساس روابط ۴ تا ۹ تعریف می‌شوند

$$\text{homogeneity} = \sum_i^{N-1} \sum_j^{N-1} \frac{P_d(i,j)}{1 + |i-j|} \quad (۴)$$

$$\text{contrast} = \sum_i^{N-1} \sum_j^{N-1} P_d(i,j) (i - j)^2 \quad (۵)$$

$$\text{dissimilarity} = \sum_i^{N-1} \sum_j^{N-1} P_d(i,j) |i - j| \quad (۶)$$

1. mean
2. variance
3. homogeneity
4. contrast
5. dissimilarity
6. entropy
7. angular second moment
8. correlation

بسیاری بین تصویر چندطیفی و تصویر تک‌رنگ باشد. طیف به دست آمده از این روش در تحقیقات پیشین توانسته است بیشترین شباهت را با طیف تصویر ابرطیفی ایجاد کند (Pande et al., 2009). دلیل دیگر مرجع قرار دادن این تحقیق و روش CNT نزدیک بودن منطقه مورد مطالعه این پژوهشگر به منطقه مورد مطالعه این تحقیق، از لحاظ پوشش‌های کاربری مشترک است. قابل ارائه بودن تحلیل‌های روش جنگل تصادفی، نظیر اهمیت باندها به علت تعداد باند کمتر در مقایسه با دیگر روش‌ها نیز، دلیل بعدی محسوب می‌شود. پس از انتخاب باندهای مشترک بین تصویر چندطیفی و تصویر تک‌رنگ و تبدیل قدرت تفکیک مکانی تصویر چندطیفی براساس قدرت تفکیک تصویر تک‌رنگ، ادغام به روش CNT طبق رابطه ۱۰ صورت می‌گیرد

$$MS_{FK}^H = \frac{SMS_K \times Pan}{\sum_k MS} \quad \text{رابطه (۱۰)}$$

در این رابطه SMS_K ، DN پیکسل در باند K ام، Pan، DN پیکسل تصویر تک‌رنگ و $\sum_k MS$ مجموع DN های همان پیکسل‌ها در همه باندهاست.

۳-۴- طبقه‌بندی تصاویر

به منظور طبقه‌بندی کاربری‌ها، چهارده کلاس عمده موجود در سطح منطقه، با بررسی میدانی و استفاده از تصاویر با قدرت تفکیک مکانی بالا انتخاب شدند و براساس آن‌ها، کلاس‌های موضوعی ارائه شده در جدول ۲ تعریف شدند. برخی کلاس‌های شناسایی شده در پیمایش‌های میدانی در اشکال ۳ و ۴ ارائه شده‌اند.

جدول ۲. کلاس‌های مورد استفاده در طبقه‌بندی

شماره کلاس	نام کلاس	تعداد نمونه‌های آموزشی	شماره کلاس	نام کلاس	تعداد نمونه‌های آموزشی
۱	مزارع ذرت	۳۴۲	۸	جاده‌ها و معابر	۴۲
۲	زمین بایر	۱۹۲	۹	ردیف درخت	۵۷
۳	زمین بایر دارای خاکروبه	۱۳۳	۱۰	باغ غیرمثمر	۴۱
۴	زمین‌های آیش	۱۲۶	۱۱	زمین‌های شخم‌خورده	۱۷۰
۵	باغ مثمر	۵۱	۱۲	صنایع کوچک و پراکنده	۱۰۹
۶	زمین خاکی بدون پوشش گیاهی	۶۹	۱۳	انبارها و تأسیسات شهری	۴۷
۷	بافت مسکونی	۲۷۵	۱۴	صنایع بزرگ و متمرکز	۱۰۳

۱. CNT

بافت عوارض مختلف، تأثیر چشمگیری داشته باشد. در این تحقیق از چهار اندازه پنجره ۳ و ۵ و ۷ و ۹، برای محاسبه ماتریس هم‌وقوعی استفاده شد. بهینه‌ترین فاصله میان زوج پیکسل‌ها در تحقیقات پیشین به صورت یک فاصله افقی-عمودی و یک فاصله افقی (1,1) = d تعیین شده است (Hall-Beyer, 2007) که در این تحقیق نیز، این فاصله برای محاسبه میان زوج پیکسل‌ها انتخاب شد.

در ادامه تحقیق، به علت وجود تعداد معتنابهی از پیکسل‌های نمونه‌های آموزشی و آزمایشی با مقادیر گنگ، مانند بی‌نهایت، در بردار ویژگی همبستگی که موجب حذف آن‌ها از طبقه‌بندی و حذف برخی کلاس‌ها می‌شد، این بردار ویژگی از فرایند طبقه‌بندی کنار گذاشته شد.

۳-۳- ادغام تصاویر

ادغام تصاویر فرایند ترکیب دو یا چند تصویر مختلف برای تولید تصویری جدید است که با استفاده از الگوریتم معینی صورت می‌پذیرد (Dobhal, 2008). ادغام تصاویر امکان دارد بر حسب مرحله رخ دادن ادغام، در سه سطح پردازشی گوناگون روی دهد:

۱. پیکسل ۲. ویژگی ۳. سطح تصمیم‌گیری

در این پژوهش تصاویر با استفاده از روش ادغام تبدیل رنگ نرمالیزه^۱ ادغام شده‌اند که از روش‌های ادغام در سطح پیکسل است. این رویکرد از روش‌های ادغام، مبتنی بر مؤلفه‌های رنگی، است که از تصویر تک‌رنگ به صورت مستقیم استفاده می‌کند. کارایی این روش در صورتی تأمین خواهد شد که همپوشانی طیفی



ردیف درخت

زمین های آیش

شکل ۳. برخی کلاس‌های استفاده‌شده در طبقه‌بندی



باغ مثمر

باغ غیر مثمر

شکل ۴. برخی کلاس‌های استفاده‌شده در طبقه‌بندی

ترتیب، ۸۰٪ کل نمونه‌ها برای آزمایش طبقه‌بندی و ۲۰٪ باقی‌مانده برای آموزش طبقه‌بندی‌کننده به کار رفتند (Gualtieri, 2009).

عملیات طبقه‌بندی با استفاده از روش جنگل‌های تصادفی انجام شده است. علت انتخاب این روش محدود بودن نمونه‌های آموزشی و عملکرد مناسب در مقایسه با دیگر الگوریتم‌های طبقه‌بندی است (Rodriguez-Galiano et al., 2012). از دیگر مزیت‌های این روش طبقه‌بندی، مصونیت در برابر انطباق بیش از حد^۲ است که نیاز به برخی عملیات، مانند اعتبارسنجی بینابینی^۳ را برطرف می‌کند (Ghimire et al., 2010; Breiman, 2001).

1. stratified
2. overfitting
3. cross-validation

روش برداشت داده‌های زمینی با توجه به وسعت کم منطقه، در دسترس بودن و همچنین، نزدیک بودن زمان گرفتن تصویر به صورت بررسی‌های میدانی بوده که در چند نوبت انجام شده است. البته پیش از بررسی‌های میدانی، با استفاده از تصاویر قدیمی‌تر با وضوح بالا از این منطقه، کاربری‌های مختلف منطقه به صورت کلی شناسایی شده بود. در بررسی‌های میدانی نیز، با استفاده از دستگاه GPS قابل حمل، موقعیت مکانی و کاربری‌های گوناگون شناسایی و به‌منزله نمونه‌های آموزشی به الگوریتم طبقه‌بندی‌کننده معرفی شد. نمونه‌های آزمایشی و آموزشی براساس انتخاب تصادفی از میان کل نمونه‌ها به نسبت ۴ به ۱ و براساس روش انتخاب تصادفی لایه‌ای^۱ تولید شدند که تعداد نمونه‌های آموزشی در جدول ۲ آمده است. بدین

درخت‌ها به وسیله کاربر که به ترتیب با ntree و mtry مشخص می‌شوند، درخت‌های تصمیم‌گیری رشد می‌یابند. رشد درخت‌ها تا رسیدن به معیار مورد نظر ادامه می‌یابد که در این روش، همان معیار کاهش بیشینه ناخالصی جینی است. در انتها نیز، براساس رأی بیشینه تمام درخت‌ها، نوع کلاس پیکسل مورد نظر انتخاب می‌شود. نمونه‌های خارج از کیسه نیز برای تعیین دقت درخت‌های تصمیم‌گیری و میزان اهمیت متغیرهای مختلف در طبقه‌بندی به کار می‌رود (Liaw et al., 2002; Richards, 2013).

برای طبقه‌بندی ابتدا با جستجوی شبکه‌ای^۶ تعریف‌شده در هر پنج سناریو، پارامترهای ntree و mtry با کمترین میزان خطای OOB شناسایی و سپس، بهترین ترکیب پارامترهای ntree و mtry وارد طبقه‌بندی شدند.

در گام بعد، برای تعیین اهمیت بردارهای ویژگی از قابلیت تشخیص باندهای مهم در روش طبقه‌بندی جنگل‌های تصادفی استفاده شد. معیارهای mean decrease Gini و mean decrease accuracy جهت سنجش اهمیت متغیرها در نرم‌افزار R مورد استفاده قرار می‌گیرد. برای محاسبه این دو معیار، نخست طبقه‌بندی‌کننده بررسی متغیر مورد نظر را متوقف می‌کند، در حالی که باقی متغیرها ثابت نگاه داشته می‌شوند. سپس، با استفاده از تخمین خطای OOB میزان کاهش دقت محاسبه می‌شود که مقدار نرمال‌شده طبق تعداد درخت‌ها در معیار mean decrease accuracy متغیر مورد نظر درج می‌شود. مجموع همه کاهش‌های شاخص ناخالصی جینی مربوط به متغیر مورد نظر نیز در معیار mean decrease Gini بازتاب می‌یابد. متغیرهای با بیشترین میزان کاهش دقت یا ناخالصی جینی مهم‌تر انگاشته می‌شوند (Kuhn et al., 2008).

1. classification and regression tree (CART)
2. bagging
3. Gini impurity
4. node
5. out-of-bag (OOB)
6. grid search

جنگل‌های تصادفی را نخستین بار برایمن در سال ۲۰۰۱ به جامعه علمی معرفی کرد (Breiman, 2001). این روش شامل مجموعه‌ای از درخت‌های تصمیم‌گیری و رگرسیون^۱ است که براساس تکنیک بگینگ^۲ اقدام به طبقه‌بندی می‌کنند (Breiman, 1996). در روش درخت تصمیم‌گیری و رگرسیون، از درختی واحد برای تقسیم فضای بین نمونه‌های آموزشی، طبق تصمیم‌گیری دوتایی، استفاده می‌شود. تقسیم‌بندی براساس معیارهایی مانند بیشترین کاهش انترופی یا کاهش بیشینه معیار ناخالصی جینی^۳ صورت می‌گیرد که به ترتیب، در روابط ۱۱ و ۱۲ آمده‌اند (Richards, 2013). تقسیم‌بندی بر طبق یکی از این معیارها و تا رسیدن به کمینه آن ادامه می‌یابد.

$$\text{رابطه (۱۱)} \quad i(N) = -\sum_j P(w_j) \log P(w_j)$$

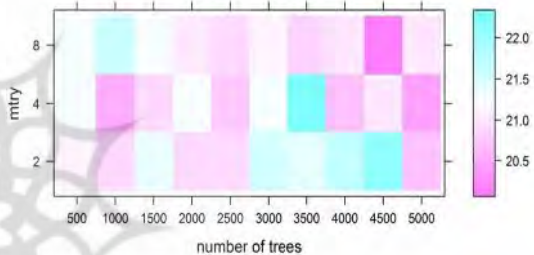
$$\text{رابطه (۱۲)} \quad i(N) = -\sum_j P(w_j) \log P(w_j)$$

در این روابط، $P(w_j)$ درصد پیکسل‌های آموزشی در گره N درخت است که در کلاس w_j واقع شده‌اند. تکنیک بگینگ بر مبنای ساخت چندین درخت و با استفاده از انتخاب تصادفی T با جایگذاری از میان نمونه‌های آموزشی انجام می‌گیرد. در انتها، با میانگین‌گیری از درخت‌ها نتیجه نهایی حاصل می‌شود. استفاده از چندین درخت برای تصمیم‌گیری باعث کاهش واریانس، ثبات بیشتر نتایج و نیز بهبود دقت می‌شود. انتخاب تصادفی نمونه‌ها با جایگذاری باعث انتخاب نشدن برخی نمونه‌ها می‌شود که در اصطلاح به آن‌ها نمونه‌های خارج از کیسه^۴ می‌گوییم. این نمونه‌ها در انتها، برای تعیین دقت طبقه‌بندی‌کننده به کار می‌روند (Liaw et al., 2002; Breiman, 1996).

جنگل‌های تصادفی کارکردی مشابه تکنیک بگینگ دارند. با این تفاوت که در این روش، به جای استفاده از همه متغیرها در فرایند تقسیم، تعدادی از آن‌ها به صورت تصادفی به کار می‌روند. سپس، طبق پارامترهای تعریف‌شده تعداد درخت و تعداد متغیرهای تصادفی از میان تمام متغیرهای مورد استفاده در تقسیم‌بندی

۴- بحث و نتایج

طبقه‌بندی براساس پنج سناریوی تعریف‌شده در بخش‌های پیشین جهت بررسی اثر اندازه پنجره و بردارهای ویژگی روی دقت طبقه‌بندی انجام شد. در سناریوی نخست، به بررسی کارایی باندهای طیفی تصویر ادغام‌شده به روش CNT پرداختیم. شکل ۵ مقادیر بهینه پارامترهای $mtry$ و $ntrre$ را در سناریوی یک نشان می‌دهد. بهترین ترکیب طبق معیار پایین ترین خطای OOB در تعداد درخت ۴۵۰۰ و مقدار پارامتر $mtry$ برابر ۸ به دست آمد. طبقه‌بندی به روش جنگل‌های تصادفی برطبق این مقادیر صورت پذیرفت.



شکل ۵. جستجوی شبکه‌ای در تصویر ادغامی به روش CNT

در طبقه‌بندی به روش جنگل‌های تصادفی از نمونه‌های آزمایشی خارج‌مانده از طبقه‌بندی برای سنجش دقت طبقه‌بندی درخت‌های تولیدی استفاده می‌شود. نتایج این طبقه‌بندی در شکل ۶ نمایش داده شده است. در این طبقه‌بندی، خطای بسیار بالایی به ترتیب صعودی در کلاس‌های باغ غیرمثمر، صنایع بزرگ و متمرکز، صنایع کوچک و پراکنده، جاده‌ها و معابر، انبارها و تأسیسات شهری مشاهده شد که در کلاس انبارها و تأسیسات شهری این خطا به ۰.۹ (۹۰٪) رسید. علت این میزان بالای خطا را می‌شود در وجود پیکسل‌های مخلوط زیاد در این کلاس‌ها و همچنین، پایین بودن دقت مکانی تصویر حاصل از روش CNT برای بررسی این کاربری‌ها دانست.

جدول ۳ نتایج طبقه‌بندی را برای کلاس‌های مختلف نشان می‌دهد. بررسی‌ها بیان می‌کند باندهای

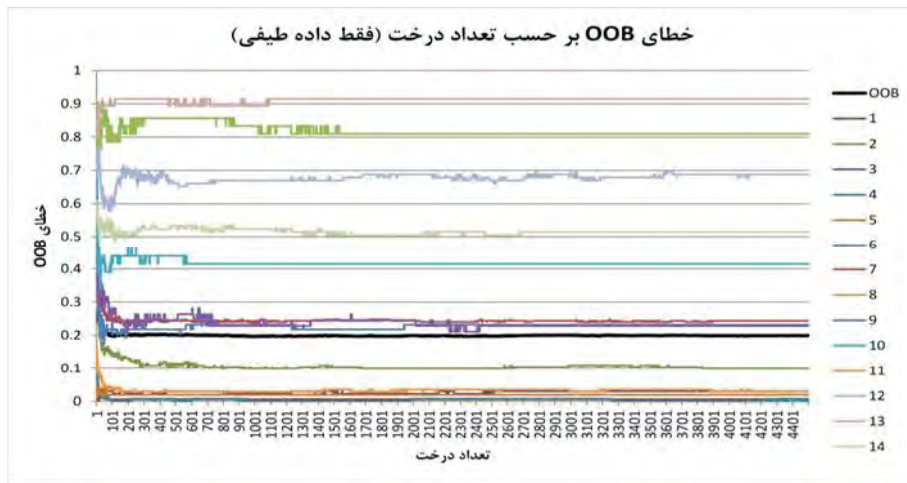
طیفی در طبقه‌بندی پوشش‌های گیاهی و پوشش‌های زمینی، مانند زمین‌های آیش و شخم‌زده و بایر، عملکرد مثبتی داشته‌اند. عملکرد باندهای طیفی در پوشش‌های شهری بیشتر ضعیف بوده است و بی‌تعادلی بالایی میان دقت تولیدکننده و دقت مصرف‌کننده در کلاس‌هایی مانند انبارها و تأسیسات شهری و جاده‌ها و معابر به چشم می‌آید. در کلاس جاده‌ها و معابر، این امر را می‌شود ناشی از دقت نامناسب تصویر ادغامی برای طبقه‌بندی این عارضه مکانی و همچنین، اختلاط بالای این کلاس به علت عرض پایین و خطی بودنشان دانست. کلاس انبارها و تأسیسات شهری نیز از علتی مشابه، یعنی تنوع زیاد مصالح تشکیل‌دهنده و ناکافی بودن طبقه‌بندی‌کننده‌های پیکسل‌منا و لزوم استفاده از طبقه‌بندی‌کننده‌های درون‌پیکسلی برای طبقه‌بندی این پوشش‌ها دانست. علت اختلاف بسیار میان دقت طبقه‌بندی کلاس‌های باغ‌های مثمر و غیرمثمر را نیز می‌شود در اختلاط طیف خاک و گیاهان بین ردیف درختان در کلاس باغ غیرمثمر به شمار آورد. این در حالی است که باغ مثمر تاج پوشش بزرگی دارد و در نتیجه، انعکاس طیفی یکنواخت‌تری را ایجاد می‌کند.

جدول ۳. دقت کلاس‌های گوناگون در طبقه‌بندی به روش

جنگل‌های تصادفی در تصویر ادغامی به روش CNT

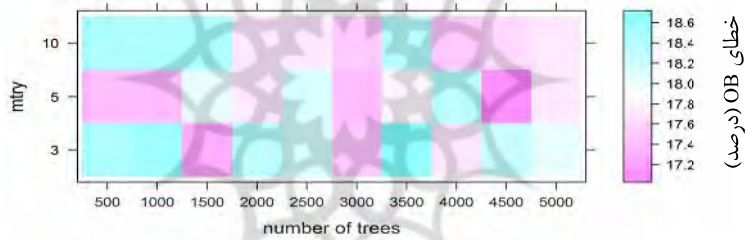
نام کلاس	دقت	
	تولیدکننده	مصرف‌کننده
مزارع ذرت	۹۹.۲۱	۹۹.۷۴
زمین بایر	۹۰.۹۹	۷۴.۳۱
زمین بایر دارای خاکروبه	۹۹.۲۵	۹۸.۸۸
زمین‌های آیش	۹۹.۲۱	۹۳.۳۱
باغ مثمر	۹۸.۰۶	۱۰۰
زمین خاکی بدون پوشش گیاهی	۸۴.۰۲	۶۸.۹۱
بافت مسکونی	۷۴.۴۲	۶۲.۸۱
جاده‌ها و معابر	۲۱.۱۵	۵۰
ردیف درخت	۶۳.۹۱	۷۴.۶۲
باغ غیرمثمر	۷۲.۲۲	۵۸.۷۹
زمین‌های شخم‌خورده	۹۷.۳۵	۹۰.۳۱
صنایع کوچک و پراکنده	۳۷.۴۴	۵۷.۹۵
انبارها و تأسیسات شهری	۶.۴۵	۳۳.۳۳
صنایع بزرگ و متمرکز	۵۲.۱۸	۶۸.۰۴

دقت کل = ۸۰.۴۷٪، ضریب کاپا = ۰.۷۸۱۱

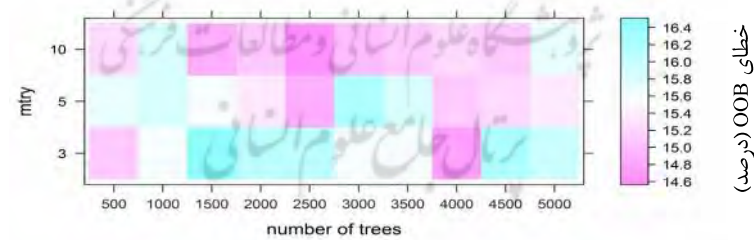


شکل ۶. خطای OOB بر حسب تعداد درخت در تصویر ادغامی به روش CNT

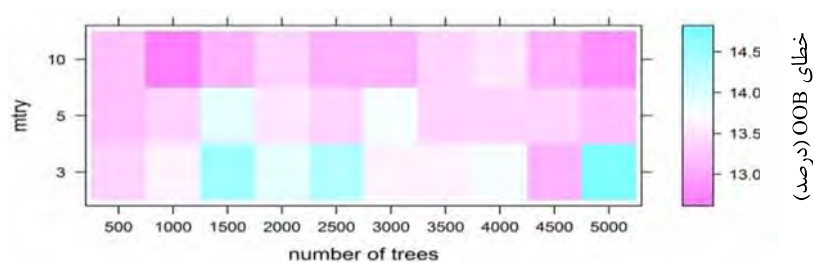
چهار سناریوی بعدی براساس افزودن بردارهای ویژگی با چهار اندازه پنجره (۳، ۵، ۷ و ۹) طراحی شدند. مراحل طی شده مشابه طبقه‌بندی باندهای طیفی بود و افزایش اندازه پنجره سبب کاهش خطای OOB شد، به گونه‌ای که خطای OOB از ۲۰٪ به کمتر از ۹٪ در اندازه پنجره ۹ کاهش یافت (شکل‌های ۱۰-۷).



شکل ۷. جستجوی شبکه‌ای در تصویر ادغامی به روش CNT و بافت با اندازه پنجره ۳

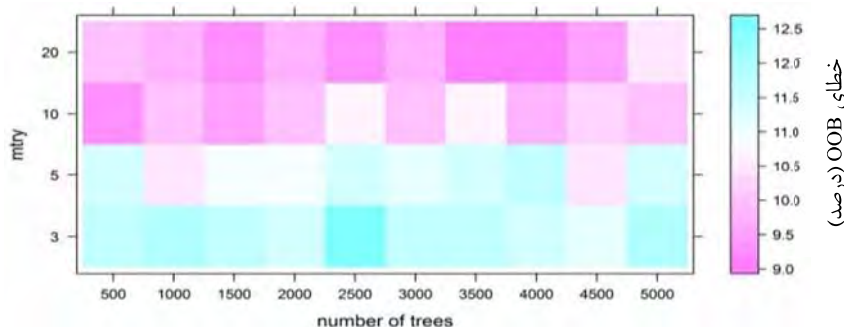


شکل ۸. جستجوی شبکه‌ای در تصویر ادغامی به روش CNT و بافت با اندازه پنجره ۵



شکل ۹. جستجوی شبکه‌ای در تصویر ادغامی به روش CNT و بافت با اندازه پنجره ۷

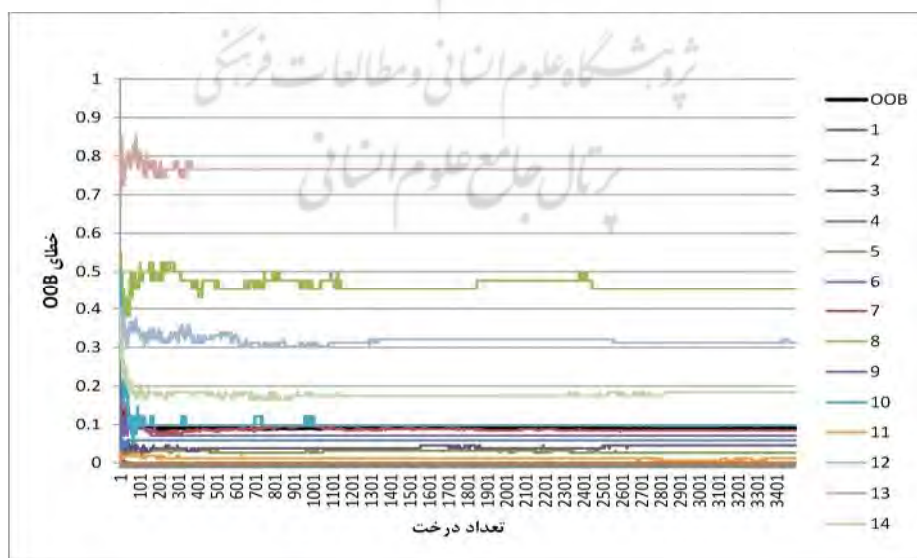
بررسی بهبود دقت طبقه‌بندی با استفاده از ادغام تصویر تک‌باند ALI ...



شکل ۱۰. جستجوی شبکه‌ای در تصویر ادغامی به روش CNT و بافت با اندازه پنجره ۹

۳۲٪، صنایع کوچک و پراکنده حدود ۳۷٪، جاده‌ها و معابر حدود ۳۶٪ و در کلاس انبارها و تأسیسات شهری این میزان کاهش خطا حدود ۱۳٪ است. این کاهش اثربخشی بافت را در بهبود دقت طبقه‌بندی نشان می‌دهد. البته همچنان در کلاس انبارها و تأسیسات شهری میزان خطا بالاست که نیاز به داده، با قدرت تفکیک مکانی بالاتر، را می‌رساند. در ادامه، طبقه‌بندی برای این چهار سناریو انجام شد که مقادیر دقت به دست آمده تولیدکننده و مصرف‌کننده در جدول ۴ ارائه شده است.

روند کاهش خطای OOB در هر چهار سناریو مشاهده شد و با افزایش اندازه پنجره محاسبه بافت، خطای OOB هر کلاس نیز کاهش بیشتری پیدا کرد. برای نمونه، در شکل ۱۱ که خطای OOB را بر حسب تعداد درخت در اندازه پنجره ۹ نشان می‌دهد، پنج کلاس مورد بررسی دارای خطای بالای OOB در سناریوی باندهای طیفی یعنی باغ غیرمثمر، صنایع بزرگ و متمرکز، صنایع کوچک و پراکنده، جاده‌ها و معابر و انبارها و تأسیسات شهری کاهش چشمگیری را شاهد بودند. کاهش خطای OOB در کلاس باغ غیرمثمر بیش از ۳۰٪، صنایع بزرگ و متمرکز بیش از



شکل ۱۱. خطای OOB بر حسب تعداد درخت، بر اثر افزودن بافت با اندازه پنجره ۹ به تصویر ادغامی به روش CNT

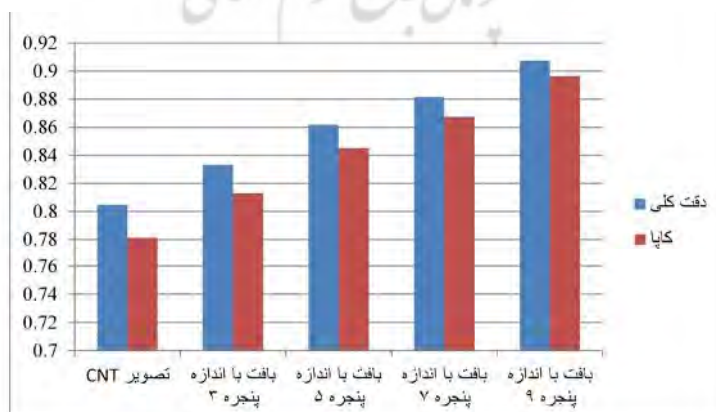
جدول ۴. دقت طبقه‌بندی بر اثر افزودن بافت با اندازه‌های مختلف

نام کلاس	اندازه پنجره ۳		اندازه پنجره ۵		اندازه پنجره ۷		اندازه پنجره ۹	
	د	م	د	م	د	م	د	م
مزارع ذرت	۹۹.۱۳	۹۹.۸۲	۹۹.۱۳	۹۹.۵۶	۹۹.۲۱	۹۹.۸۲	۹۹.۳۰	۹۹.۸۲
زمین بایر	۹۳.۰۸	۸۱.۴۹	۹۷.۳۹	۸۷.۴۶	۹۸.۴۳	۹۲.۴۰	۹۹.۰۹	۹۶.۹۳
زمین بایر دارای خاکروبه	۹۷.۳۷	۹۹.۶۲	۹۶.۹۹	۹۹.۸۱	۹۶.۲۴	۱۰۰	۹۶.۹۹	۹۹.۰۴
زمین‌های آیش	۹۹.۶۰	۹۷.۶۷	۹۹.۶۰	۹۷.۸۶	۹۹.۶۰	۹۸.۰۵	۹۹.۶۰	۹۷.۸۶
باغ مثمر	۹۸.۵۴	۱۰۰	۹۶.۶۰	۱۰۰	۹۶.۶۰	۱۰۰	۹۷.۵۷	۱۰۰
زمین خاکی بدون پوشش گیاهی	۸۹.۰۴	۸۹.۴۵	۹۰.۴۱	۹۰.۴۱	۹۴.۹۸	۹۴.۹۸	۹۲.۲۴	۹۷.۱۲
بافت مسکونی	۸۱.۷۱	۶۳.۶۴	۸۵.۸۰	۶۹.۱۸	۸۷.۰۶	۷۱.۵۴	۹۱.۸۳	۷۸.۹۳
جاده‌ها و معابر	۲۶.۹۲	۵۲.۵۰	۲۹.۴۹	۷۶.۶۷	۴۵.۵۱	۷۵.۵۳	۵۴.۴۹	۷۱.۴۳
ردیف درخت	۸۰.۴۳	۷۷.۷۳	۸۶.۹۶	۸۲.۶۴	۸۴.۳۵	۸۶.۶۱	۸۱.۳۰	۸۳.۴۸
باغ غیرمثمر	۷۹.۶۳	۸۰.۱۲	۸۰.۲۵	۸۳.۳۳	۸۸.۲۷	۸۰.۳۴	۸۷.۰۴	۷۶.۲۲
زمین‌های شخم‌خورده	۹۶.۹۱	۹۷.۰۵	۹۶.۹۱	۹۷.۳۴	۹۷.۹۴	۹۸.۸۱	۹۸.۶۸	۹۸.۳۹
صنایع کوچک و پراکنده	۴۶.۵۸	۵۲.۴۴	۵۶.۸۵	۶۰	۶۰.۹۶	۶۷.۰۹	۷۴.۲۰	۷۷.۰۱
انبارها و تأسیسات شهری	۵.۹۱	۶۴.۷۱	۷.۵۳	۷۰	۱۱.۸۳	۷۵.۸۶	۱۷.۷۴	۷۸.۵۷
صنایع بزرگ و متمرکز	۵۲.۴۳	۶۹.۰۱	۶۴.۸۱	۷۰.۶۳	۷۴.۰۳	۷۲.۶۲	۸۳.۵۰	۸۰.۱۹
دقت کلی طبقه‌بندی	۸۳.۳۲		۸۶.۱۸		۸۸.۱۵		۹۰.۷۴	
مقدار کاپا	۰.۸۱		۰.۸۴۵۲		۰.۸۶۷۳		۰.۸۹۶۵	

اندازه پنجره بر دقت طبقه‌بندی، می‌شود از معیار کاپای شرطی استفاده کرد که امکان مقایسه میان کلاس‌ها را در حالت‌های گوناگون فراهم می‌کند (جدول ۵). بیشترین مقادیر کاپای شرطی در هر کلاس با حالت

با افزایش اندازه پنجره، دقت تولیدکننده «د ت» و مصرف‌کننده «د م» در بیشتر کلاس‌ها افزایش می‌یابد. برای مقایسه بهتر روند افزایش دقت، مقادیر دقت کلی و کاپا در شکل ۱۲ آمده است.

اگرچه دقت تولیدکننده و مصرف‌کننده روندی افزایشی را نشان می‌دهند، برای سنجش دقیق‌تر اثر



شکل ۱۲. نمودار دقت کلی و کاپا در سناریوهای مختلف

جدول ۵. مقادیر کاپای شرطی بر اثر افزودن بافت با اندازه‌های مختلف

نام کلاس	صفاً باندهای طیفی	بافت با اندازه پنجره ۳	بافت با اندازه پنجره ۵	بافت با اندازه پنجره ۷	بافت با اندازه پنجره ۹
مزارع ذرت	۰.۹۹۶۸	۰.۹۹۷۹	۰.۹۹۴۷	۰.۹۹۷۹	۰.۹۹۷۹
زمین بایر	۰.۷۰۹۷	۰.۷۹۰۸	۰.۸۵۸۳	۰.۹۱۴۲	۰.۹۶۵۴
زمین بایر دارای خاکروب	۰.۹۸۷۸	۰.۹۹۵۸	۰.۹۹۷۹	۱	۰.۹۸۹۶
زمین‌های آیش	۰.۹۲۷۶	۰.۹۷۴۸	۰.۹۷۶۹	۰.۹۷۸۹	۰.۹۷۶۹
باغ مثمر	۱	۱	۱	۱	۱
زمین خاکی بدون پوشش گیاهی	۰.۶۷۸۶	۰.۸۹۰۹	۰.۹۰۰۹	۰.۹۸۴۱	۰.۹۷۰۲
بافت مسکونی	۰.۵۶۰۳	۰.۵۷۰۱	۰.۶۳۵۶	۰.۶۶۳۵	۰.۷۵۰۹
جاده‌ها و معابر	۰.۴۸۸۰	۰.۵۱۳۶	۰.۷۶۱۱	۰.۷۴۹۵	۰.۷۰۷۴
ردیف درخت	۰.۷۳۷۱	۰.۷۶۹۴	۰.۸۲۰۲	۰.۸۶۱۳	۰.۸۲۸۹
باغ غیرمثمر	۰.۵۷۷۷	۰.۷۹۶۳	۰.۸۲۹۲	۰.۷۹۸۵	۰.۷۵۶۲
زمین‌های شخم‌خورده	۰.۸۹۲۱	۰.۹۶۷۲	۰.۹۷۰۴	۰.۹۸۶۸	۰.۹۸۲۰
صنایع کوچک و پراکنده	۰.۵۴۹۹	۰.۴۹۱۰	۰.۵۷۱۹	۰.۶۴۷۷	۰.۷۵۴۰
انبارها و تأسیسات شهری	۰.۳۱۴۲	۰.۶۳۶۹	۰.۶۹۱۴	۰.۷۵۱۷	۰.۷۷۹۶
صنایع بزرگ و متمرکز	۰.۶۵۹۳	۰.۶۶۹۷	۰.۶۸۷۰	۰.۷۰۸۲	۰.۷۸۸۸

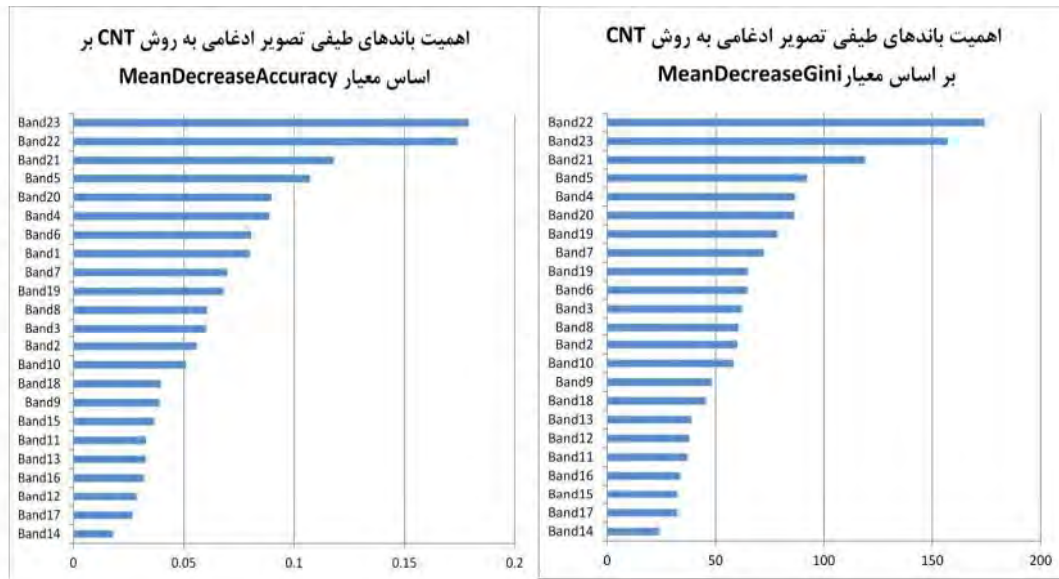
روش طبقه‌بندی جنگل‌های تصادفی استفاده شد. شکل زیر اهمیت باندهای طیفی در تصویر ادغام‌شده به روش CNT براساس این دو معیار را نشان می‌دهد. باندهای طیف قرمز تصویر CNT مانند باند ۲۱، ۲۲ و ۲۳ در هر دو تصویر بهترین عملکرد را داشتند و ضعیف‌ترین عملکرد به باندهای واقع در طیف سبز متعلق بود. علت این امر را می‌شود تفکیک‌پذیری مناسب‌تر، دامنه گسترده‌تر مقادیر DN و همپوشانی کمتر در بین کلاس‌های مختلف در این بازه طیفی دانست.

در سناریوهای گوناگون افزودن بافت نیز این دو معیار به کار رفت. نتایج بررسی این دو معیار برای اندازه پنجره ۹ نیز ارائه شده که بالاترین دقت کلی را در طبقه‌بندی به دست آورد (شکل ۱۴).

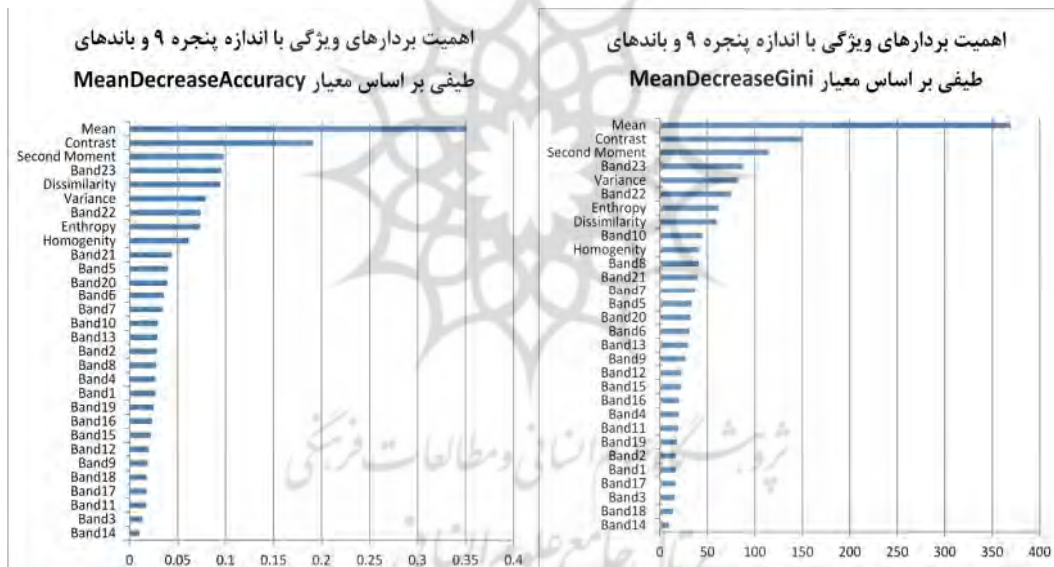
نمودارهای شکل ۱۵ براساس میانگین‌گیری از مقادیر نمونه‌های آزمایشی هر کلاس و در ادامه نرمال کردن آن به دست آمده است. بردار ویژگی میانگین که بهترین عملکرد را داشت، دارای بازه گسترده‌ای در کلاس‌های گوناگون بود و تقریباً هیچ‌گونه همپوشانی بین مقادیر آن مشاهده نمی‌شد. در بردار ویژگی

به غیر از کلاس باغ مثمر که در آن، تمامی پنجره حالت عملکرد یکسانی دارند، در سایر کلاس‌ها برتری با حالت‌های افزودن بافت است. این برتری، به‌ویژه در برخی کلاس‌ها مانند انبارها و تأسیسات شهری و صنایع کوچک و پراکنده مشهود است. بهبود عملکرد در کلاس‌های مرتبط با مناطق شهری مانند صنایع بزرگ و متمرکز، انبارها و تأسیسات شهری، بافت مسکونی و صنایع کوچک و پراکنده با افزایش اندازه پنجره مشاهده می‌شود. دلیل این امر را می‌شود گسترده شدن دامنه متغیرهای ماتریس هم‌وقوعی در اندازه پنجره ۹، به نسبت دیگر اندازه‌ها دانست که ممکن است پیچیدگی بافت این مناطق را که از نظر طیفی مصالح گوناگون مانند آجر، بتن، سیمان و آسفالت دارند، بهتر منعکس کند. در دیگر کلاس‌ها مانند زمین‌های خاکی، مزارع ذرت و باغ مثمر که بافت همگن‌تری دارند، استفاده از بافت با اندازه‌های بالاتر الزاماً سبب افزایش دقت نمی‌شود.

از دو معیار mean decrease accuracy و mean decrease Gini برای تعیین اهمیت بردارهای ویژگی در



شکل ۱۳. اهمیت باندهای مختلف براساس معیارهای موجود در تصویر ادغامی به روش CNT



شکل ۱۴. اهمیت باندهای مختلف در حالت افزودن بافت با اندازه پنجره ۹ به تصویر ادغامی به روش CNT

و در نتیجه این بردار، در مقایسه با دو بردار ویژگی قبل، ویژگی عملکرد نامناسب تری داشت. در انتها، نقشه موضوعی طبقه بندی برای سناریوهای مختلف تهیه شد که بهترین عملکرد در بین پنج سناریو، یعنی سناریوی شماره ۵ (افزودن بافت با اندازه پنجره ۹ به تصویر ادغامی به روش CNT) در تصویر (ب) با نقشه موضوعی طبقه بندی تصویر ادغامی (بدون افزودن بافت) (سناریوی ۱) تصویر (الف) و

کنتراست، میان دو کلاس زمین های شخم خورده و صنایع کوچک و پراکنده تداخل وجود داشت که شناسایی و تمایز میان این دو کلاس را در این بردار ویژگی مشکل کرد. وضعیت تداخل میان کلاس ها در بردار ویژگی ممان دوم زاویه ای شدیدتر بود و پنج کلاس انبارها و تأسیسات شهری، صنایع بزرگ و متمرکز، ردیف درخت، جاده ها و معابر و بافت مسکونی مقادیر مشابه داشتند که تمایز میان آن ها را سخت کرد

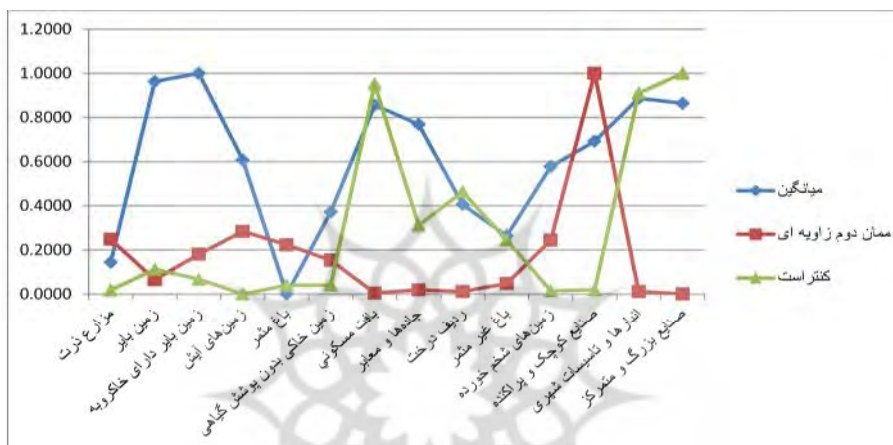
تفکیک مکانی بالا و تصویر ابرطیفی سنجنده هایپریون امکان مناسبی را برای استفاده از اطلاعات مکانی، مانند بافت این تصویر تک‌رنگ، فراهم می‌آورد. استفاده از بافت این تصویر تک‌رنگ در بهبود دقت طبقه‌بندی تصویر ادغامی از سنجنده هایپریون در این پژوهش هدف‌گذاری شد و مورد مطالعه قرار گرفت.

الگوریتم جنگل‌های تصادفی روشی کارآ و ساده بود که با تنظیم فقط دو پارامتر می‌شود آن را پیاده‌سازی

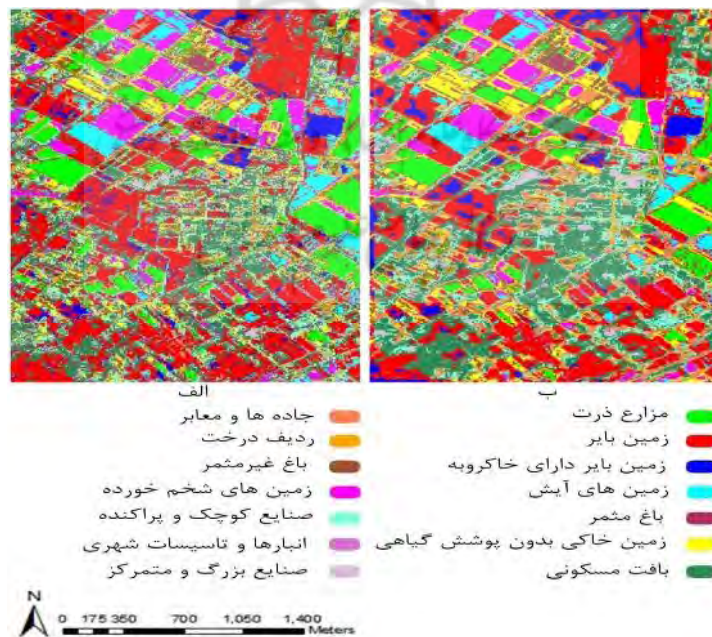
در شکل ۱۶ آمده است. بهبود لبه‌ها، حذف نویز موجود در نقشه مربوط به تصویر CNT، حذف پیکسل‌های غیرمرتبط، بارسازی بهتر مناطقی مانند بافت مسکونی و جاده‌ها و معابر و چنین مواردی را می‌شود از مزایای نقشه تولیدشده به وسیله افزودن بافت دانست.

۵- نتیجه‌گیری

گرفتن همزمان تصویر تک‌رنگ سنجنده ALI با قدرت



شکل ۱۵. مقادیر میانگین نرمال شده کلاس‌های گوناگون سه بردار ویژگی برتر



شکل ۱۶- مقایسه نقشه طبقه‌بندی تصویر بر اثر افزودن بافت و تصویر تنها با باندهای طیفی

۶- سپاسگزاری

در پایان، بر خود لازم می‌دانیم تا از سازمان USGS برای در اختیار گذاشتن رایگان داده‌ها و همۀ توسعه‌دهندگان نرم‌افزار R تشکر کنیم.

۷- منابع

- Breiman, L., 1996, **Bagging Predictors**, Machine learning, 24(2): 123-140.
- Breiman, L., 2001, **Random Forests**, Machine learning, 45(1): 5-32.
- Chutia, D., Bhattacharyya, D.K., Kalita, R., Goswami, J., Singh, P.S. & Sudhakar, S., 2014, **A Model on Achieving Higher Performance in the Classification of Hyperspectral Satellite Data: A Case Study on Hyperion Data**, Applied Geomatics, 6(3): 181-195.
- Dobhal, S., 2008, **Performance Analysis of High-Resolution and Hyperspectral Data Fusion for Classification and Linear Feature Extraction**, Msc Thesis, ITC, Enschede.
- Fawagreh, K., Gaber, M.M. & Elyan, E., 2014, **Random Forests: From Early Developments to Recent Advancements**, Systems Science & Control Engineering: An Open Access Journal, 2:1-21.
- Ghimire, B., Rogan, J. & Miller, J., 2010, **Contextual Land-Cover Classification: Incorporating Spatial Dependence in Land-Cover Classification Models Using Random Forests and the Getis Statistic**, Remote Sensing Letters, 1(1): 45-54.
- Gualtieri, J.A., 2009, **The Support Vector Machine (SVM) Algorithm for Supervised Classification of Hyperspectral Remote Sensing Data**, In: Kernel Methods for Remote Sensing Data Analysis, Edited by Camps-Valls, G., & Bruzzone, L., Wiley, New York, pp. 51-83.
- Hall-Beyer, M., 2007, **GLCM Texture: A Tutorial, Version 2010**, Available in: <http://www.fp.ucalgary.ca/mhallbey/tutorial.htm>.

کرد. این روش در این تحقیق نیز توانست کارآیی مناسبی از خود نشان دهد. بررسی خروجی‌ها و نتایج حاصل از این روش طبقه‌بندی نشان داد که استفاده از بافت در طبقه‌بندی ممکن است پوشش‌های کاربری شهری با پیکسل‌های دارای اختلاط بالای طیفی مانند ردیف درخت، جاده‌ها و معابر و بافت مسکونی و انواع کاربری‌های صنعتی را بهبود چشمگیری بخشد. با افزودن داده‌های بافت استخراج‌شده از تصویر تک‌رنگ سنجنده ALI به تصویر ابرطیفی، دقت کلی طبقه‌بندی در حدود ۱۰٪ افزایش یافت و دقت کلاس‌های گوناگون تا ۳۰٪ رشد را نشان داد. همچنین خطای OOB که از پارامترهای مهم در طبقه‌بندی جنگل‌های تصادفی است، حدود ۱۱٪ کاهش داشت. افزایش اندازه پنجره در محاسبه ویژگی‌های بافت به افزایش دقت کلی طبقه‌بندی منجر شد. بیشترین دقت نیز هنگام افزودن بافت با اندازه پنجره ۹ به داده‌های طیفی حاصل شد. یکی از خروجی‌های مهم الگوریتم جنگل تصادفی، تشخیص باندها و متغیرهای مهم در طبقه‌بندی است. این ویژگی طیف متنوعی از گزینه‌های تصمیم‌گیری را براساس اهمیت باندها در اختیار کاربر قرار می‌دهد که برای حذف باندهای با کمترین اهمیت و کاهش ابعاد داده کارآیی دارد. برخی بردارهای ویژگی ماتریس هم‌وقوعی مانند میانگین، کنتراست و ممان دوم زاویه‌ای، و باندهای طیفی محدوده قرمز مانند باند ۲۲، ۲۱ و ۲۳ از نظر اهمیت در طبقه‌بندی جنگل‌های تصادفی، رتبه‌های بالا را به دست آوردند. با توجه به کاربرد گسترده تصاویر ابرطیفی در زمینه‌هایی مانند مدیریت جنگل‌ها، شناسایی کانی‌ها و طبقه‌بندی پوشش‌های اراضی، پیشنهاد می‌شود رویکرد افزودن بافت تصاویر تک‌رنگ به آن‌ها مورد بررسی قرار گیرد. نیز، می‌شود دیگر روش‌های ادغام تصاویر را مانند GST، PCT برای طبقه‌بندی به کار برد و اثر آن‌ها در تغییرات دقت طبقه‌بندی بررسی کرد.

- Resolution Panchromatic Imagery for Urban Land-Use Classification**, Remote Sensing of Environment, 113(6): 129-1276.
- Pande, H., Tiwari, P.S. & Dobhal, S., 2009, **Analyzing Hyper-Spectral and Multi-Spectral Data Fusion in Spectral Domain**, Journal of the Indian Society of Remote Sensing, 37:395-408.
- Pour, A.B. & Hashim, M., 2013, **Fusing ASTER, ALI and Hyperion Data for Enhanced Mineral Mapping**, International Journal of Image and Data Fusion, 4:1-20.
- Qiu, F., 2008, **Neuro-Fuzzy Based Analysis of Hyperspectral Imagery**, Photogrammetric Engineering & Remote Sensing, 74(10): 1235-1247.
- Richards, J.A., 2013, **Remote Sensing Digital Image Analysis: An Introduction**, Springer, Verlag Berlin Heidelberg, pp. 413-421.
- 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, 67:93-104.
- Welikanna, D.R., Tolpekin, V. & Kant, Yogesh, 2008, **Analysis of the Effectiveness of Spectral Mixture Analysis and Markov Random Field Based Super Resolution Mapping Over an Urban Environment**, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, 38: 641-649.
- Zhang, H., Zhang, Y. & Lin, H., 2012, **Urban Land Cover Mapping Using Random Forest Combined with Optical and SAR Data**, IEEE International In Geoscience and Remote Sensing Symposium (IGARSS): 6809- 6812.
- Haralick, R.M., Shanmugam, K., Dinstein, I.H., 1973, **Textural Features for Image Classification**, IEEE Transactions on Systems, Man and Cybernetics, 6: 610-621.
- Haralick, R.M., 1979, **Statistical and Structural Approaches to Texture**, Proceedings of the IEEE, 67(5): 786-804.
- Hsu, P.H., 2007, **Feature Extraction of Hyperspectral Images Using Wavelet and Matching Pursuit**, ISPRS Journal of Photogrammetry and Remote Sensing, 62(2): 78-92.
- Kasetkasem, T., Arora, M. K. & Varshney, P.K., 2005, **Super Resolution Land Cover Mapping Using a Markov Random Field Based Approach**, Remote Sensing of Environment, 96: 302-314.
- Kuhn, S., Egert, B., Neumann, S. & Steinbeck, C., 2008, **Building blocks for Automated Elucidation of Metabolites: Machine Learning Methods for NMR Prediction**, BMC Bioinformatics, 9(1): 400-412.
- Liaw, A., & Wiener, M., 2002, **Classification and Regression by Random Forest**, R news, 2(3).
- Licciardi, G.A., Khan, M.M. & Chanussot, J., 2012, **Fusion of Hyperspectral and Panchromatic Images: A Hybrid Use of Indusion and Nonlinear PCA**, 19th IEEE International Conference on Image Processing (ICIP): 2133-2136.
- Lu, D. & Weng, Q., 2007, **A Survey of Image Classification Methods and Techniques for Improving Classification Performance**, International journal of Remote sensing, 28(5): 823-870.
- Murray, H., Lucieer, A. & Williams, R., 2010, **Texture-Based Classification of Sub-Antarctic Vegetation Communities on Heard Island**, International Journal of Applied Earth Observation and Geoinformation, 12: 138-149.
- Pacifici, F., Chini, M. & Emery, W. J., 2009, **A Neural Network Approach Using Multi-Scale Textural Metrics from Very High-**