ارزیابی کارایی سه الگوریتم ماشین بردار پشتیبان، درخت تصمیم گیری و شبکه عصبی مصنوعی آرتمپ فازی در تهیه نقشه اراضی جنگلی (مطالعه موردی: جنگل‌های شمال استان ایلام)

صالح آرخی[1]، علی اکبر جعفرزاده[2]، محمد ملکی[3]

تاریخ دریافت: 90/2/5 تاریخ پذیرش: 91/7/22

چکیده

طبقه‌بندی پوشش زمین با استفاده از تصاویر ماهواره‌ای یکی از مهم‌ترین کاربردهای سنجش از دور است و الگوریتم‌های زیادی برای این منظور توسعه یافته‌اند. هدف از این تحقیق ارزیابی کارایی سه الگوریتم ماشین‌بردار پشتیبان، درخت تصمیم-گیری و شبکه عصبی مصنوعی فازی آرتمپ جهت تهیه نقشه اراضی جنگلی شمال استان ایلام با استفاده از تصویر سنجنده **LissIII** مربوط به سال 2007 می‌باشد. در این تحقیق پس از تصحیح هندسی تصاویر و پیش‌پردازش‌های صورت‌گرفته، اقدام به تعیین نمونه‌های تعلیمی مربوط به طبقات پوشش‌زمین برای انجام عملیات طبقه‌بندی شد و ارزیابی تفکیک‌پذیری نمونه‌ها در هر کدام از طبقات پوشش‌زمین با استفاده از شاخص آماری واگرایی تعیین شد. در مرحله بعد، به‌منظور ارزیابی صحت نتایج طبقه‌بندی هر کدام از الگوریتم‌ها، نقشه واقعیت زمینی با استفاده از روش سیستماتیک به ابعاد 550 متری در منطقه مورد مطالعه طراحی و نوع پوشش زمینی آن‌ها در محل نمونه‌ها مشخص گردید. در نهایت، کارایی هر کدام از روش-های طبقه‌بندی با استفاده از معیارهای صحت کلی، ضریب کاپا، صحت تولیدکننده و صحت کاربر مورد بررسی قرار گرفت. مقایسه صحت‌های کلی و ضریب کاپای کسب شده برای سه طبقه‌بندی کننده با مجموعه باندی مناسب در مقایسه با نقشه واقعیت زمینی بیانگر این بود که طبقه‌بندی کننده ماشین‌بردار پشتیبان با مقدار صحت کلی 94/36 درصد و ضریب کاپای 90/21 درصد نتایج مطلوب‌تری نسبت به سایر الگوریتم‌ها دارد. نتایج نشان‌داد که تفکیک و طبقه‌بندی اراضی جنگلی از سایر طبقات کاربری منطقه با صحت بالا صورت گرفته‌است.

واژه‌های کلیدی: اراضی جنگلی، سنجنده **LissIII**، ارزیابی صحت، طبقه‌بندی نظارت‌شده، جنگل زاگرس

1- استادیار گروه جغرافیا، دانشکده علوم انسانی، دانشگاه گلستان، نویسنده مسئول

saleh148@ yahoo.com

2- کارشناس ارشد جنگلداری، دانشکده منابع طبیعی دانشگاه علوم کشاورزی و منابع طبیعی ساری

jafarzadeh82@ yahoo.com

3- کارشناس ارشد عمران- سنجش از دور، گروه عمران، دانشکده فنی ومهندسی دانشگاه ایلام

mohammad_maleky2014@ yahoo.com

مقدمه

در بسیاری از فعالیت‌های مدیریتی و برنامه‌ریزی، داشتن اطلاعاتی درباره پوشش زمینی و کاربری اراضی حایز اهمیت بسیاری بوده و برای مدل‌سازی و شناخت زمین به‌عنوان یک سیستم، امری ضروری محسوب می‌گردد. اصطلاح پوشش‌زمینی[1] به انواع پدیده‌های موجود در سطح زمین اطلاق می‌گردد (لیلساند و کیفر[2]، 1997). اطلاعات مربوط به پوشش‌زمینی را می‌توان به‌طور مستقیم از تصاویر ماهواره‌ای مناسب استخراج نمود. یکی از تحلیل‌های مهمی که بر روی تصاویر ماهواره‌ای انجام می‌گیرد، طبقه‌بندی این تصاویر می‌باشد. از آن‌جا که هدف اصلی از پردازش تصاویر ماهواره‌ای، تهیه نقشه-های موضوعی و کارآمد می‌باشد، انتخاب الگوریتم مناسب طبقه‌بندی نقش زیادی در این خصوص ایفا می‌کند. انواع مختلفی از الگوریتم‌های طبقه‌بندی از زمان کسب اولین تصاویر ماهواره‌ای توسعه یافته‌اند. روش‌های طبقه‌بندی با نظارت به دو گروه پارامتریک و غیر پارامتریک تقسیم می‌شوند. مشکل اصلی روش-های طبقه‌بندی پارامتریک (مثل حداکثر احتمال، حداقل فاصله و غیره)، وابستگی آن‌ها به توزیع آماری داده‌ها (مثل توزیع نرمال گوسن) می‌باشد. به‌همین دلیل روش‌های غیر پارامتریک، از جمله روش‌های مبتنی بر شبکه‌های عصبی مصنوعی[3] و درخت تصمیم‌گیری[4] و ماشین بردار پشتیبان[5] برای طبقه‌بندی تصاویر ماهواره‌ای به‌کار گرفته شده‌است.

اگرچه روش‌های شبکه‌عصبی مصنوعی ممکن است با دقت بالا جهت طبقه‌بندی تصاویر ماهواره‌ای مورد استفاده قرار گیرد، اما تعدادی از مطالعات نشان داده‌اند که استفاده‌کنندگان از روش‌های شبکه‌عصبی مشکلاتی در زمینه انتخاب پارامترهای گوناگون در طول اجرای یادگیری دارند. انتخاب معماری شبکه، اندازه نمونه برای آموزش، یادگیری الگوریتم‌ها و تعداد تکرارهای مورد نیاز برای آموزش برخی از مشکلات این روش می‌باشند (واپنیک،[6] 1995).

الگوریتم طبقه‌بندی دیگر، درخت تصمیم-گیری است که از یک رویکرد متفاوت جهت طبقه‌بندی تصاویر ماهواره‌ای استفاده می‌کند. این روش، اغلب مشکل پیچیدگی طبقه‌بندی را به مراحل چندگانه با فرآیندهای تصمیم‌گیری ساده-تر تبدیل می‌نماید. روش‌های درخت تصمیم-گیری مزیت‌های برای طبقه‌بندی تصاویر ماهواره‌ای بخاطر انعطاف‌پذیری، ماهیت غیر پارامتریکی و توانایی اجرا روابط غیر پارامتریکی بین عوارض و کلاس‌ها دارند. دقت یک طبقه-بندی انجام‌شده به‌وسیله روش‌های درخت تصمیم‌گیری بستگی به فاکتورهای از قبیل اندازه و ترکیب داده‌های تعلیمی، روش انتخاب مشخصه، انتخاب روش هرس کردن و نوع روش طبقه‌بندی درخت تصمیم‌گیری دارد که از جمله معایب آن محسوب می‌شود.

[1] - Land Cover
2 -Lilliesand and kiefer
[3]- Artificial Neural Network
[4] - Decision Tree
5 -Support Vector Machine

6 -vapnik,1995

در چند سال اخیر یک تکنیک جدید براساس تئوری یادگیری آماری با نام ماشین‌های بردار پشتیبان (SVM) جهت طبقه‌بندی داده‌های سنجش از دور مورد استفاده واقع شده‌است (دیکسون و همکاران [1]،2008). ماشین‌های بردار پشتیبان الگوریتم‌های سنجش نظارت شده براساس تئوری سنجش آماری هستند.

در زمینه طبقه‌بندی پوشش‌زمینی با استفاده از داده‌های سنجش از دور تحقیقات متعددی وجود دارد. اولین طبقه‌بندی پوشش سطح زمین در مقیاس جهانی با روش حداکثر احتمال توسط دفریس و توانشید [2](1994) و تهیه نقشه پوشش سطح زمین با قدرت تفکیک 1 کیلومتر با استفاده از رویکرد طبقه‌بندی نظارت نشده توسط لولند و همکاران (2000) انجام گرفته‌است. به‌علت محدودیت‌های این روش‌ها، در سال‌های اخیر، رویکردهای طبقه‌بندی درختی و شبکه عصب مصنوعی که خصوصیات غیرخطی و غیر پارامتریکی دارند، در سطح منطقه‌ای و جهانی مورد استفاده قرار گرفتند. هانسن و همکاران [3](1996) از داده‌های NOAA/AVHR برای تهیه نقشه پوشش سطح زمین در مقیاس جهانی و روش‌های حداکثر احتمال و طبقه‌بندی درختی استفاده کرده‌اند. در این بررسی دقت روش طبقه-بندی درختی نسبت به حداکثر احتمال بهتر گزارش شده‌است. شبکه عصبی روش دیگری می‌باشد که به‌طور گسترده در طبقه‌بندی داده‌های سنجش از دور مورد بررسی واقع شده‌است. در شناسایی محدوده گسترش گندم، مورتی و همکاران [4](2003) از این روش استفاده کردند که کارآمدی بهتری نسبت به روش حداکثر احتمال نشان‌داد. برای غلبه به مشکل ساختاری شبکه عصبی، روش آرتمپ فازی [5] توسعه یافت که این روش می‌تواند منطق فازی و تئوری تشدید انطباقی را با هم تلفیق کند.گوپال و همکاران [6](1999) با استفاده از رویکرد آرتمپ فازی پوشش سطح زمین را در مقیاس جهانی طبقه-بندی نموده‌اند. بوراک و همکاران [7](1999) روش‌های حداکثر احتمال، درخت تصمیم‌گیری و شبکه عصب مصنوعی آرتمپ فازی را مقایسه نموده و نتیجه گرفتند که روش طبقه‌بندی درختی می‌تواند بیشتر اطلاعات تصویر را حفظ نموده و نسبت به نمونه تعلیمی حساس نمی‌باشد در حالیکه آرتمپ فازی نسبت به درخت تصمیم-گیری و حداکثر احتمال دقت بالاتری دارد. ماشین بردار پشتیبان یکی دیگر از روش‌های طبقه بندی تصاویر ماهواره‌ای است که در این تحقیق مورد بررسی قرار گرفته‌است. روش ماشین‌های بردار پشتیبان یک روش کلاسه‌بندی الگو است که اولین بار توسط وپنیک [8] معرفی شد. دیکسون و کن داد [9](2008)، از الگوریتم ماشین‌های بردار پشتیبان برای طبقه‌بندی تصاویر ماهواره‌ای استفاده کردند و نتایج این طبقه‌بندی کننده را با الگوریتم شبکه عصبی مقایسه کردند و نتیجه

1 -Dixon and Yao et al,2008
2 -Defries and Townshed
3 -Hanson and et al
4 -Murthy and et al
5 -Fuzzy Artmap
6 -Gupal and et al,1999
7 -Borak and et al,1999
8 -Vapink
9 -Dixon ans Candade

گرفتند که الگوریتم ماشین‌های بـردار پـشتیبان دارای دقت و صحت بالای در امر طبقه‌بندی می-باشد.

تحقیقاتی که در داخل کشور صورت گرفتـه است با استفاده از روش‌های معمول طبقه‌بنـدی مثل الگوریتم حداکثر احتمال، حداقل فاصله از میانگین و متوازی السطوح بوده‌است (شتایی و عبدی، 1386). تا به حال هیچ مطالعـه جـامع و کاملی با استفاده از روش‌های مورد اشاره واقـع شده در این مطالعه انجام نشده‌است، لذا در این مطالعه از الگوریتم‌های ماشین‌های بردار پشتیبان، درختی و شبکه عصب مصنوعی آرتمـپ فـازی برای طبقه‌بندی پوشش اراضی کوهستانی استان ایلام و با استفاده از تصویر ماهواره‌ای LissIII کسب شده در سال 2007 اسـتفاده شـده‌اسـت. همچنین کارآمدی این روش‌ها با هم مقایـسه و بصورت آماری آنالیز گردیده‌است.

روش‌های طبقه‌بندی

روش شبکه عصبی مصنوعی فازی آرتمپ

روش شبکه عصبی مصنوعی آرتمپ فازی از روش‌های جدید در امر طبقه‌بندی تصاویر ماهواره‌ای است و به دو صورت نظارت‌شده و نظارت‌نشده می‌باشد. در شبکه نظارت‌شده، ورودی به سیستم اعمال و اختلاف بین خروجی شبکه و خروجی مطلوب برای تغییر و تطبیق وزن‌ها به‌کار می‌رود. در صورتی‌که در شبکه نظارت‌نشده، شبکه خود را متناسب با داده‌های ورودی تطبیق می‌کند. در این مطالعه، روش نظارت‌شده مورد استفاده قرار گرفت.

روش طبقه‌بندی درخت تصمیم‌گیری

طبقه‌بندی درختی مورد استفاده در این مطالعه، یک روش تک‌متغیره غیرپارامتریک می‌باشد. در عمل، این الگوریتم مشخصه‌ای (از قبیل باند انعکاسی) را به‌صورت تکراری انتخاب نموده، نمونه‌هایی که قابلیت تفکیک دارند را به دو گروه تقسیم نموده و تفاوت در داخل هر زیر گروه را به حداقل می‌رساند ضمن این‌که بین گروه‌ها حداکثر تمایز را ایجاد می‌کند.

برای رسم درخت طبقه‌بندی، کاربر ابتدا بایستی از داده‌های تعلیمی استفاده کند که به این مرحله، مرحله یادگیری داده‌های تعلیمی می-گویند. سپس کل تصویر بر اساس همین داده‌های تعلیمی طبقه‌بندی می‌شود. برای شروع کار، تمام پیکسل‌های تعلیمی از تمام کلاس‌ها به ریشه اختصاص می‌یابند. از آن‌جایی‌که ریشه حاوی تمام پیکسل‌های تعلیمی از همه کلاس‌ها است، یک فرایند تکراری برای رشد درخت آغاز و کلاس‌ها را از یگدیگر جدا می‌سازد (میلر و فرانکلین، 2001[1]).

الگوریتم‌های ماشین بردار پشتیبان

الگوریتم SVM[2]، بر پایه تئوری یادگیری آماری که توسط واپنیک[3]، ابداع شد، بنا شده است. براساس این تئوری می‌توان کران نرخ خطای ماشین یادگیری را برای داده‌های طبقه-بندی نشده، به‌عنوان نرخ خطای تعمیم‌یافته، در نظر گرفت. این کران‌ها به‌عنوان تابعی از مجموع نرخ خطاهای آموزشی هستند که میزان پیچیدگی

1 - Miller & Franklin, 2001
2- Support Vector Machine
3 - Vapnik

طبقه‌بندی کننده‌ها را نشان می‌دهند. به‌منظور کمینه‌کردن نرخ خطاهای تعمیم یافته، بایستی نرخ خطای آموزش و پیچیدگی طبقه‌بندی کننده کاهش یابد. این کار را می‌توان با بیشینه‌کردن حاشیه جداسازی انجام داد .

مواد و روش‌ها

منطقه مورد مطالعه

منطقه مورد مطالعه در این تحقیق، یک حوزه کوهستانی از مناطق جنگلی شمال استان ایلام می‌باشد. مساحت این منطقه 225593 هکتار بوده، که از لحاظ موقعیت جغرافیایی، بین طول‌های شمالی "32 '19 °33 تا "07 '60 °33 و عرض‌های شرقی "48 '00 °46 تا "43 '00 °47 واقع شده‌است (شکل2). ارتفاع منطقه حداقل 700 متر و حداکثر 2700 متر بالاتر از سطح دریا می‌باشد (نقشه DEM منطقه). از لحاظ شرایط اقلیمی، این منطقه جزو مناطق نیمه خشک (اقلیم مدیترانه‌ای) با متوسط بارندگی 592/78 میلی‌متر در سال می‌باشد (اداره هواشناسی استان ایلام). اراضی جنگلی منطقه پوشیده از جنگل‌های بلوط- بنه و بلوط ایرانی با تراکم‌های مختلف، اراضی مرتعی، زمین‌های زراعی دیم و آبی می‌باشد. کلاس کاربری پوشش زمینی انتخاب‌شده در این مطالعه و تعریف آن‌ها به‌شرح زیر است:

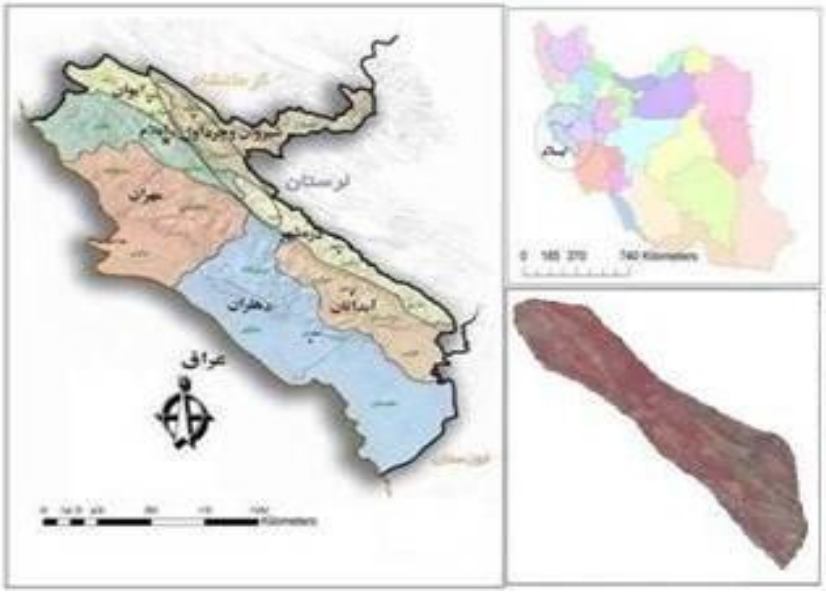
کلاس مرتع: شامل اراضی با مرتع مشجر و غیر مشجر، اراضی دارای پوشش گیاهی خودرو و فاقد جنگل

کلاس جنگل: به مناطق جنگلی گفته می‌شود که از طرف سازمان جنگل‌ها و مراتع تعریف شده است.

کلاس مناطق مسکونی: کلیه مراکز انسان ساخت اعم از شهرها، روستاها، مراکز صنعتی

کلاس اراضی کشاورزی: کلیه اراضی زیر کشت دیم و آبی

کلاس اراضی لخت: شامل مناطق سنگلاخی و بدون پوشش



شکل1- موقعیت کشوری و استانی منطقه مورد مطالعه

روش تحقیق

در این تحقیق، ازتصویر ماهواره‌ای سنجنده LissIII مربوط به سال 2007 منطقه، استفاده شد. به‌منظور بررسی وجود و یا عدم وجود خطاهای هندسی و رادیومتری، بررسی کیفیت داده‌ها بر روی تصاویرماهواره‌ای صورت گرفت. با بررسی تصاویر تک باند و ترکیبات رنگی مختلف هیچ‌گونه خطای قابل ملاحظه‌ای مشاهده نشد. این تصاویر با استفاده از بسته نرم افزاری ENVI4.5 و به کار بردن تبدیل چندجمله‌ای، به سیستم تصویر UTM در ناحیه 38 شمالی تبدیل شد. برای تصحیح هندسی از نقاط کنترل زمینی به دست‌آمده از روی نقشه‌های توپوگرافی 1:50000، عکس‌های هوایی 1:20000 و نقاط مرجع زمینی با استفاده از GPS استفاده شد. همچنین عمل نمونه‌گیری مجدد با استفاده از روش نزدیکترین همسایه[1] به‌منظور جلوگیری از تغییر ارزش‌های طیفی تصاویر به‌کار گرفته شد.

پس از آن که داده‌ها مورد تصحیح هندسی قرار گرفتند، برای بهره‌مندی از توان اطلاعاتی داده‌ها، به کمک الگوریتم‌های طبقه‌بندی، قابلیت آن‌ها برای تفکیک کلاسه‌های موضوعی که در تحقیق مورد نظر هستند، مورد بررسی و آزمون قرار گرفتند. براساس تجربه و نتایج تحقیقات گذشته، باندهای مصنوعی نظیر مؤلفه‌های اصلی و نسبت‌گیری‌هایی مثل NDVI و RVI به همراه باندهای اصلی سنجنده LissIII می‌توانند برای تفکیک بهتر کلاس‌ها مورد استفاده قرار گیرند. ایجاد این باندها خود نوعی عملیات بارزسازی بوده که در نتیجه منجر به تغییر DNs[2] شده و جزو عملیات قبل از طبقه‌بندی محسوب می‌شوند. شاخص‌های مزبور پوشش گیاهی با استفاده از تصاویر سنجنده LissIII تولید شدند. در این تحقیق از روش تجزیه مؤلفه اصلی به روش استاندارد نیز استفاده شد (ساورویی، 1378). در نهایت پس از مشاهده آماره‌های مربوط به واریانس، میانگین و همچنین هیستوگرام پراکنش مؤلفه‌های اصلی ایجاد شده از بین آن‌ها مؤلفه‌های اول و دوم را که دارای بیشترین واریانس بودند جهت استفاده در ترکیبات باندی برای عمل طبقه‌بندی مورد استفاده قرار گرفتند (لیلساند و کیفر3، 1997).

طبقه‌بندی و تهیه نقشه پوشش زمینی

با هدف تفکیک پوشش‌های زمینی عمده منطقه بر روی داده‌های ماهواره‌ای تطابق هندسی شده، از روش‌های طبقه‌بندی نظارت‌شده و سه الگوریتم ماشین‌های بردار پشتیبان، شبکه عصبی مصنوعی آرتمپ فازی و روش درخت تصمیم‌گیری استفاده شد. انتخاب نمونه‌های تعلیمی مورد نیاز برای طبقه‌بندی کلاس‌های مختلف در منطقه با شناخت از منطقه و با استفاده از اطلاعات جنبی و استفاده از تصویر رنگی مرکب 4،3،2 (RGB) صورت گرفت. انتخاب این تصویر رنگی مرکب صرفاً جهت شناسایی کلاس‌های مورد نظر و انتخاب بهترین نمونه‌های تعلیمی بوده‌است. انتخاب نمونه‌های تعلیمی یکی

[1] - Nearest Neighbor

[2] - Digital Numbers

3 -Lilliesand and kiefer

از مهمترین فاکتورهای تعیین‌کننده درجه تعمیم قواعد طبقه‌بندی به نمونه‌های نامشخص می‌باشد. برای این منظور در منطقه مورد مطالعه به فراخور سهم هر طبقه تعداد مناسبی نمونه تعلیمی به‌طور تصادفی با استفاده از بررسی‌های میدانی، تصاویر گوگل ارث[1] و تصاویر رنگی مرکب انتخاب گردیدند. پس از آن‌که نمونه‌های تعلیمی برای پوشش‌های مختلف انتخاب شدند، جهت ارزیابی و اصلاح نمونه‌ها، نشانه‌های طیفی هر یک از کلاس‌های فوق در باندهای اصلی و پردازش‌شده استخراج شد و نمودار توزیع ارزش طیفی نمونه-های تعلیمی در تمامی باندهای طیفی بررسی گردید. برای بررسی تشابه آن‌ها و میزان تفکیک-پذیری و تباین کلاس‌ها، از روش ارزیابی کمی تفکیک‌پذیری استفاده و تفکیک‌پذیری آن‌ها با استفاده از شاخص واگرایی بررسی شد. با هدف انتخاب بهترین مجموعه باندهای که تفکیک-پذیری طبقات برای طبقه‌بندی در آن‌ها بهتر صورت پذیرد از مشخصه‌های آماری نمونه‌های تعلیمی استفاده گردید که مبتنی بر واگرایی بین نشانه‌های طیفی مربوط به طبقات بر پایه محاسبه میانگین و ماتریس واریانس - کوواریانس طبقات در نمونه‌های تعلیمی انتخابی می‌باشند. پس از انتخاب بهترین ترکیبات باندی عملیات طبقه‌بندی با استفاده از الگوریتم های ماشین بردار پشتیبان، شبکه عصبی مصنوعی آرتمپ فازی و روش درختی صورت گرفت. به‌منظور حذف پیکسل-های منفرد و پراکنده در سطح تصویر طبقه‌بندی شده و همچنین به‌دست آوردن تصویر مطلوب و با وضوح بیشتر، از فیلتر پایین‌گذر نما در اندازه 3x3 پیکسل استفاده شد.

تهیه نقشه واقعیت زمینی

در این تحقیق برای تعیین صحت نقشه‌های حاصل از الگوریتم‌های مختلف طبقه‌بندی داده-های ماهواره‌ای، با استفاده از روش نمونه‌برداری تصادفی سیستماتیک، یک شبکه نمونه‌برداری 550 متری طراحی و بر روی نقشه منطقه پیاده شد و با در نظر گرفتن موقعیت مکانی محل قطعات نمونه در روی زمین، نوع پوشش سطح زمین تعیین شد. در روی زمین موقعیت نمونه‌ها با استفاده از دستگاه موقعیت‌یاب جهانی (GPS) با دقت بالا پیاده و نوع پوشش زمینی در محل نمونه‌ها مشخص گردید. باتوجه به اطلاعات به دست آمده از عملیات صحرایی و با بهره‌گیری از قابلیت سامانه‌های اطلاعات جغرافیایی نقشه واقعیت زمینی با پنج کلاس تهیه گردید. به‌دلیل ساختار رستری تصاویر ماهواره‌ای و نیز با توجه به این‌که این نقشه مبنا در ارزیابی دقت طبقه‌بندی تصاویر مورد استفاده قرار می‌گیرد، ساختار نقشه واقعیت زمینی از حالت برداری به رستری تبدیل شد. در مجموع 495 نقطه از منطقه مورد مطالعه برداشت و از آن‌ها یک نقشه با ساختار رستری تهیه و در نهایت نقشه‌های حاصل از الگوریتم-های مختلف طبقه‌بندی با نقشه واقعیت زمینی مقایسه شده و ماتریس خطا[2]، تشکیل گردید.

[1] - Google Earth

[2] - Confuse Matrix

روش‌های ارزیابی الگوریتم‌های طبقه‌بندی

معیارهای بررسی عملکرد الگوریتم‌های طبقه‌بندی یا ارزیابی آن‌ها شامل: دقت، سرعت، پایداری و جامع‌بودن می‌باشد. این‌که کدام معیار یا کدام گروه از معیارها استفاده شوند، بستگی به هدف مطالعه دارد. معیاری که به تمام اهداف مربوط می‌باشد دقّت یا صحت است که در این بررسی به عنوان معیار ارزیابی الگوریتم‌های مختلف انتخاب شد. سرعت و پایداری همچنین فاکتورهای مهّم در ارزیابی کارایی الگوریتم‌ها هستند که در این بررسی مورد توجه قرار گرفتند. دو نوع از معیارهای پر استفاده تعیین دقّت، یعنی صحت کلی[1] و ضریب کاپا[2] در این مطالعه مورد استفاده قرار گرفتند. صحت کلی مزیت تفسیر به صورت مستقیم را دارد، در حالی که ضریب کاپا امکان آزمایش آماری تفاوت قابل ملاحظه یا همان معنی‌داری بین دو الگوریتم را فراهم می‌کند (کنگالتون[3]، 1991).

نتایج

تصحیح هندسی تصاویر با میزان خطای ریشه میانگین مربعات(RMSE) حدود 0/37 پیکسل ثبت شد و با روی هم‌گذاری لایه‌های خطی رودخانه‌ها بر روی تصویر تطابق یافته، دقت بالای عمل تطابق هندسی مورد تایید قرار گرفت.

پس از تصحیح هندسی تصاویر و اطمینان از دقت زمین مرجع بودن تصاویر، از باندهای مصنوعی نظیر مؤلفه‌های اصلی و نسبت‌گیری‌های رایج مثل NDVI و RVI، به همراه باندهای اصلی سنجنده LissIII می‌توانند برای تفکیک بهتر کلاسه‌ها مورد استفاده قرار گیرند. از بین مؤلفه‌های ایجادشده ناشی از تجزیه مؤلفه‌های اصلی، مولفه‌های اول و دوم، به عنوان مناسب‌ترین مؤلفه‌ها انتخاب و در طبقه‌بندی شرکت داده شدند. با وجود ترکیبات مختلف باندی حاصل از آنالیزهای مختلف، نتایج نشان‌دهنده مناسب بودن باندهای اصلی سنجنده LissIII (باندهای 2،3،4،5)، برای طبقه‌بندی کاربری می‌باشد.

جهت طبقه‌بندی تصویرماهواه‌ای LissIII، کلاس‌های پوشش زمینی در پنج گروه تحت عنوان کلاس‌های اراضی کشاورزی، مرتع، جنگل، اراضی بایر و مناطق مسکونی تعیین و سپس نمونه‌های تعلیمی بصورت تصادفی از سطح منطقه با استفاده از تصاویر ماهواره‌ای گوگل ارث و بازدید میدانی جمع‌آوری شد

در مرحله بعد با استفاده از ویژگی‌های تصاویر، کلاس‌های پوشش اراضی در محدوده مورد مطالعه وارد شد و تفکیک‌پذیری کلاس‌ها با معیار فاصله واگرایی تبدیل شده و با استفاده از خصوصیات آماری نمونه‌های تعلیمی مورد بررسی قرار گرفت (جدول1).

پس از مشخص نمودن میزان تفکیک‌پذیری کلاس‌ها، نسبت به طبقه‌بندی تصاویر سنجنده LissIII در بهترین ترکیب باندی به سه روش ماشین بردار پشتیبان (SVM)، شبکه عصبی مصنوعی آرتمپ فازی (FA) و روش درختی (DT) اقدام شد. بدین ترتیب نقشه‌های پوشش اراضی مربوط به سال 2007 به‌دست آمد (اشکال

1 - Overall Accuracy
2 - Kappa Coefficient
3 - Congalton, 1991

4، 5 و6). نتایج حاصل از ارزیابی معیارهای مختلف صحت در جدول 2 و همچنین نتایج حاصل از طبقه‌بندی با استفاده از سه الگوریتم ماشین‌بردار پشتیبان، شبکه عصبی مصنوعی فازی آرتمپ و روش درختی با استفاده از مجموعه باندی مناسب در جدول 3 ارایه شده‌است.

به‌طور کلی، تفاوت‌های مطلق دقت طبقه‌بندی در میان سه روش طبقه‌بندی خیلی زیاد نیست. با این حال بسیاری از تفاوت‌ها از نظر آماری معنی‌دار هستند. از نظر ثبات الگوریتم، روش SVM نسبت به سایر الگوریتم‌ها، پایدارتر می‌باشد. از دو الگوریتم دیگر، روش DT نسبت به روش FA به‌صورت محسوسی پایدارتر بوده‌است. همچنین از نظر سرعت اجرا، الگوریتم DT نسبت به دو الگوریتم دیگر سریع‌تر بوده و روش FA کم‌ترین سرعت را در اجرای عملیات طبقه‌بندی نشان‌داد.

جدول1- میزان تفکیک‌پذیری طبقات کاربری اراضی با استفاده از شاخص واگرایی.

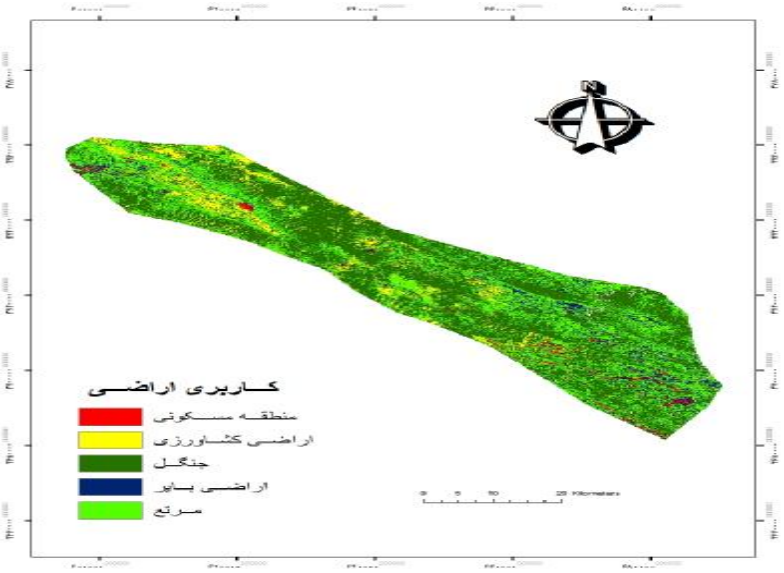
طبقات کاربری	میزان تفکیک‌پذیری	وضعیت تفکیک‌پذیری
مناطق مسکونی با اراضی کشاورزی	2/00000	خوب
مناطق مسکونی با جنگل	2/00000	خوب
مناطق مسکونی با اراضی لخت	1/73453	کم
مناطق مسکونی با مرتع	1/99864	خوب
اراضی کشاورزی با جنگل	1/98457	خوب
اراضی کشاورزی با اراضی لخت	2/00000	خوب
اراضی کشاورزی با مرتع	2/00000	خوب
جنگل با اراضی لخت	2/00000	خوب
جنگل با مرتع	1/99852	خوب
اراضی لخت با مرتع	1/64478	کم
میانگین تفکیک‌پذیری تمام طبقات	1/95074	خوب

جدول2- نتایج ارزیابی صحت روش‌های مختلف به تفکیک کلاس‌ها

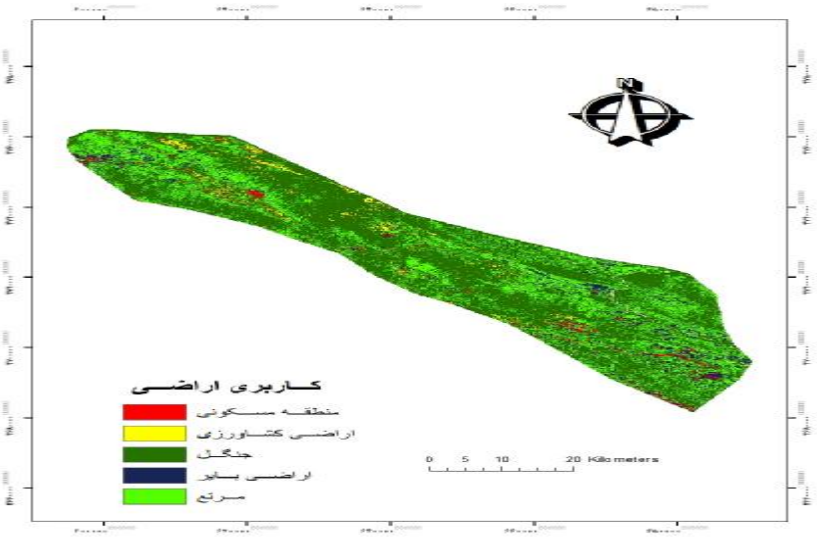
طبقات کاربری	ماشین بردار پشتیبان		درخت تصمیم‌گیری		فازی آرتمپ	
	صحت تولیدکننده	صحت کاربر	صحت تولیدکننده	صحت کاربر	صحت تولیدکننده	صحت کاربر
جنگل	97/34	92/21	96/09	90/50	90/98	85/71
اراضی کشاورزی	93/37	90/83	91/12	88/90	87	79/82
مرتع	91/78	93/65	86/30	93/50	82/15	87/69
مناطق مسکونی	96/21	94/53	93/40	94/14	86/13	88/78
اراضی لخت	94/20	91/10	92/18	87/64	84/04	77/95

جدول3- نتایج کلی ارزیابی صحت الگوریتم های طبقه‌بندی

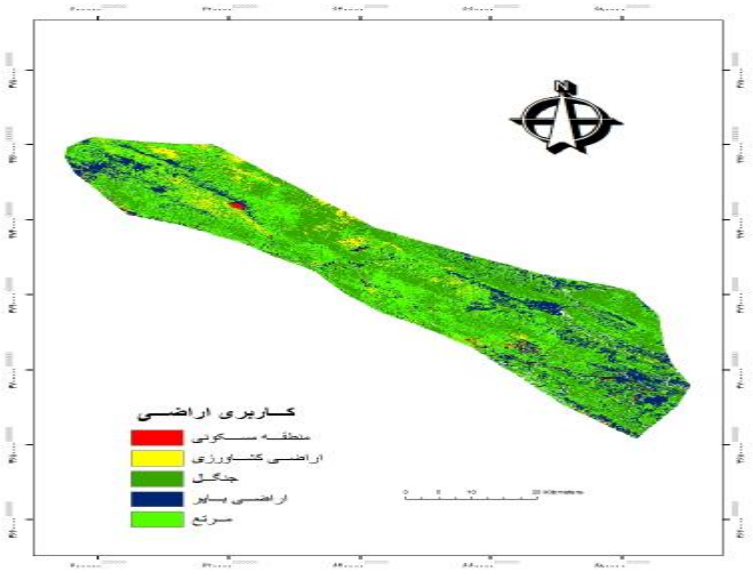
بهترین ترکیب باندی	الگوریتم طبقه بندی کننده	صحت کلی (%)	ضریب کاپا(%)
باندهای اصلی 432	ماشین بردار پشتیبان	94/36	90/21
	درخت تصمیم‌گیری	91/42	87/83
	آرتمپ فازی	88/57	80/41



شکل2- تصویر طبقه‌بندی شده با استفاده ازالگوریتم ماشین بردار پشتیبان



شکل3- تصویر طبقه‌بندی شده با استفاده از الگوریتم درخت تصمیم‌گیری



شکل4- تصویر طبقه‌بندی شده با استفاده از الگوریتم فازی آرتمپ

بحث و نتیجه‌گیری

در این مطالعه، پس از تصحیحات لازم و پیش پردازش های اولیه تصاویر، اقدام به طبقه بندی داده‌ها به سه روش مختلف گردید، به عبارت دیگر کیفیت تصاویر از نظر وجود خطاهای هندسی و رادیومتری بررسی و تصحیح گردید. هدف این مطالعه تعیین کارایی روش طبقه‌بندی ماشین بردار پشتیبان جهت طبقه‌بندی کاربری اراضی و پوشش‌گیاهی با استفاده از سری داده چندطیفی در مقایسه با روش‌های طبقه‌بندی درختی و شبکه عصبی مصنوعی فازی آرتمپ بود.

نتایج بررسی تفکیک‌پذیری طبقات و انتخاب مجموعه باندی مناسب برای طبقه‌بندی تصویر و تهیه نقشه پوشش‌زمینی منطقه نشان‌داد که با وجود استفاده از مجموعه باندهای اصلی و باندهای پردازش‌شده حاصل از آنالیزهای مختلف (باندهای حاصل از تجزیه‌وتحلیل مؤلفه اصلی و نسبت‌گیری)، باندهای اصلی سنجنده LissIII مناسب‌ترین مجموعه باندی برای تفکیک کلاس-های مورد نظر بوده‌اند که نتیجه حاصله با نتایج شتایی و عبدی (86) مطابقت دارد. نتایج بررسی تفکیک‌پذیری طبقات با معیار فاصله واگرایی تبدیل شده نشان‌داد که میزان تفکیک‌پذیری طبقات به غیر از طبقه اراضی بدون پوشش با مرتع و طبقه مناطق مسکونی با اراضی بدون پوشش که از تفکیک‌پذیری تقریباً کمی برخوردار بودند، با یکدیگر خوب بوده است. علت عدم تفکیک‌پذیری این دو طبقه را می‌توان در تشابه طیفی آن‌ها جستجو کرد. نتایج ارزیابی دقت نشان‌داد که صحت تولید طبقه جنگل در مقایسه با سایر طبقات کاربری، از صحت تولید کننده بالاتری برخوردار است.

روش طبقه‌بندی درختی برای برطرف‌سازی بسیاری از مشکلات طبقه‌بندی مورد استفاده قرار گرفته‌است. اگرچه گزارش شده‌است که این

روش کم‌وبیش نتایج طبقه‌بندی دقیق‌تری نسبت به روش‌های معمول ارایه می‌دهد، اما انتخاب بهترین روش انشعاب مبحث مهمی است که کارایی این الگوریتم را تحت تاثیر قرار می‌دهد. در این مطالعه، از روش انشعاب جینی برای طبقه‌بندی کلاس‌های کاربری اراضی منطقه مورد مطالعه استفاده شد و نتایج حاکی از کارایی بالاتر این الگوریتم نسبت به روش شبکه عصب مصنوعی فازی آرتمپ دارد. یانگ و همکاران (2003)، زو و همکاران (2005)، چوبی و همکاران (2006) و اتوکی و بلسچک (2009) در تحقیقات خود این الگوریتم طبقه‌بندی را دقیق‌ترین روش برای طبقه‌بندی تصاویر ماهواره‌ای معرفی کردند.

الگوریتم ماشین بردار پشتیبان (SVM) که اخیراً برای طبقه‌بندی تصاویر ماهواره‌ای مورد استفاده قرار گرفته‌است یکی از دقیق‌ترین روش‌ها می‌باشد که کارایی آن وابسته به انتخاب بهترین نوع کرنل(هسته) یا تابع می‌باشد. در این مطالعه از تابع(هسته) شعاعی به‌عنوان بهترین تابع کرنل برای طبقه‌بندی کلاس‌های اصلی کاربری زمین در منطقه استفاده شد.

مقادیر کاپا و صحت کلی برای هر روش طبقه‌بندی مورد استفاده در این تحقیق با داده‌های سنجنده LissIII محاسبه شد. این مطالعه نشان داد که الگوریتم ماشین‌های بردار پشتیبان با صحت کلی 94/36 درصد و ضریب کاپا 90/21 درصد دارای کارایی بیشتر و دقت بالاتری نسبت به دیگر روش‌ها می‌باشد. روش طبقه‌بندی درختی با مقدار صحت کلی 91/42 درصد و ضریب کاپا 87/83 درصد و شبکه عصب مصنوعی آرتمپ فازی با مقدار صحت کلی 88/57 درصد و ضریب کاپا 80/41 درصد در درجه اهمیت بعدی قرار دارند. این نتایج به‌طور کلی تصدیق‌کننده مطالعات پیشین می‌باشد که در آن‌ها روش SVM دقیق‌تر از روش‌های شبکه عصبی مصنوعی و درخت تصمیم‌گیری است. شفری و رامل (2009)، فودی و ماتور (2004) و هوانگ و همکاران (2002) در تحقیقات خود به این نتیجه رسیدند. همچنین این روش در تمامی کلاس‌ها نسبت به دو روش دیگر دقیق‌تر می‌باشد. علت دقت بالای الگوریتم SVM نسبت به دیگر روش‌ها را می‌توان این‌طور بیان کرد که این روش از لحاظ محاسباتی سریع بوده و همچنین این روش از قوانین بهینه‌سازی جهت مکان‌یابی مرزهای بهینه بین کلاس‌ها کاربری استفاده می‌کند در نتیجه جایگزین مناسبی برای سایر الگوریتم‌های معمول طبقه‌بندی می‌باشند.

منابع

1- ساروئی، س.، 1378. بررسی امکان طبقه‌بندی جنگل به لحاظ تراکم درجنگل‌های زاگرس به کمک داده‌های ماهواره‌ای درمنطقه سپید دشت شهرستان درود ـ لرستان. پایان‌نامه کارشناسی ارشد جنگلداری، دانشگاه تهران، صفحه110 .

2- شتایی، ش و امید عبدی.، 1386. تهیه نقشه کاربری اراضی در مناطق کوهستانی زاگرس با استفاده از داده‌های سنجنده ETM^+ (منطقه مورد مطالعه: حوزه سرخاب خرم آباد لرستان). مجله علوم کشاورزی و منابع طبیعی 14(1): 23- 13.

3- Borak, J. S. and Strahler, A.H., 1999. Feature selection and land cover classification of a MODIS-like dataset for a semiarid environment. INT. J. Remote Sensing, 20 : 919-938.

4- Carpenter, G.A., Grossberg, S. and Reynolds, J.H., 1991. ARTMAP: Suppervised Real-time learing and classification of nonstationary data by a self-organizing neural networks, Neural networks, Vol.4, pp.565-588.

5- Chubey, M.S., Franklin, S.E. and Wulder, M.A., 2006. Object-based Analysis of Ikonos-2 Imagery for Extraction of Forest Inventory Parameters. Photogrammetric Engineering & Remote Sensing 72 (4), 383-394.

6- Congalton, R., 1991. A review of assessing the accuracy of classification of remotely sensed data. Remote Sensing of Environment, 37: 35 - 46.

7- Cover, T.M., 1965.Geometrical and statistical properties of systems of linear inqualities with application in pattern recognition, IEEE Transation on Electronic Computers, vol. Ec-14, pp.326-334.

8- DeFries, R S., and Townshed, J.G., 1994. NDVI derived land cover classification at a globe scale. International Journal of Remote Sensing., 5: 3567-3586.

9- Dixon, B. and Candade, N., 2008. Multispectral land use classification using neural networks and support vector machines: one or the other, or both?, International Journal of Remote Sensing 29: 1185–1206.

10-Foody, G.M., Mathur, A., 2004. A relative evaluation of multiclass image classification by support vector machines. IEEE Transactions on Geoscience and Remote Sensing 42: 1335–1343.

11-Gopal, S., Woodcock, C.E., Strahler, A.H., 1999. Fuzzy Neural Network Classification of Global Land Cover from a 1o AVHRR DataSet. Remote Sensing of Environment, 67 : 230－243.

12--Gustavo Camps-Valls, Luis Gomez-Chova, and Javier Calpe-Maravilla., 2002. Robust Support Vector method for hyperspectral data classification and knolewdge discovery, IEEE Transaction on Geoscience and Remote sensing, vol.42,July.

13-Hanson, MC., Dubayah, R., and Defries, R., 1996. Classification trees: an alternative to traditional land cover classifers. INT. J. Remote Sensing, 17: 1075-1081.

14-Hastie, T., and Tibshirani, R., 1998. Classification by pairwise coupling. Ann. Stat. 2: 475 – 471.

15--Hsu, C.W., Chang, C.C. and Lin, C.J., 2008. A practical guide to support vector classification.http://www.csie.ntu.edu.tw/cjlin/papers/guide/guide.pdf.

16-Huang, C., Davis, L. S., and Townshend, J. R., 2002. An assessment of support vector machines for land cover classification. International Journal of Remote Sensing, 23: 725 749.

17-Lilliesand, T.M., Kiefer, R.W., 1997. Remote sensing and image interpretation. 3th ed., University of Wisconisn,Madison, 735 P.

18-Loveland, T.R., Reed, B.C., and Brown, J.F., Ohlen, D.O., Zhu, Z., Yang, L. and Merchant, J.W., 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. International Journal of Remote Sensing. 21: 1303-1330.

19-Miller, J. and Franklin, J., 2001. Modeling the distribution of four vegetation alliances using generalized linear models and classification trees with spatial dependence. Ecological Modeling, 157: 227–247.

20-Murthy, CS., Raju, P. and Bardrinath, K., 2003. Classification of wheat crop with multi-temperoral images: performance of maximum likelihood and artificial neural networks. INT. J. Remote Sensing, 24(23): 4871-4890.

21-Otukei, J.R. and Blaschke, T., 2009. Land cover assessment using decision tree, support vector machines and maximum likelihood classification algorithms.International journal of Applied Earth Observation and Geoinformation, JAG.297.pp.1-5.

22-Shafri,H. and Ramle, F., 2009. A Comparion of Support Vector Mashine and Decision Tree Classifications Using Satellite Data of Langkawi Island. Information Technology Journal, 8(1): 64 – 70.

23-Vapnik, V., 1995. The Nature of Statistical Learning Theory (New York: Springer Verlag).

24-Xu, M., Watanachaturaporn, P., Varshney ,P.K. and Arora, M., 2005. Decision tree regression for soft classification of remote sensing data. Remote Sens. Environ., 97: 322-336.

25-Yang, C., Prasher, S.O., Enright, P., Madramootoo, C., Burgess, M., Goel, P.K. and Callum , I., 2003. Application of decision tree technology for image classification using remote sensing data. Agricultural Systems 76: 1101 – 1117.

26-Zambon, M., Lawrence, R., Bunn, A. and Powell, S. 2006. Effect of alternative splitting rules on image processing using classification tree analysis. Photogrammetric Engineering and Remote Sensing, 72(1): 25–30.