



## ارزیابی الگوریتم‌های طبقه‌بندی نظارت شده غیر پارامتریک در تهیه نقشه پوشش زمین با

### استفاده از تصاویر لندست ۸

وحید میرزایی زاده<sup>۱\*</sup>، مریم نیک‌نژاد<sup>۲</sup>، جعفر اولادی قادیکلایی<sup>۳</sup>

۱. دانش‌آموخته کارشناسی ارشد جنگل‌داری، دانشگاه ایلام
۲. دانشجوی دکتری جنگل‌داری، دانشگاه علوم کشاورزی و منابع طبیعی ساری
۳. دانشیار دانشکده منابع طبیعی، دانشگاه علوم کشاورزی و منابع طبیعی ساری

#### مشخصات مقاله

پیشینه مقاله:  
 دریافت: ۶ دی ۱۳۹۳  
 پذیرش: ۲۱ تیر ۱۳۹۴  
 دسترسی اینترنتی: ۲۵ آذر ۱۳۹۴

#### واژه‌های کلیدی:

طبقه‌بندی تصویر  
 نقشه پوشش زمین  
 سنجش از دور  
 ماشین بردار پشتیبان  
 آبخیز ارکواز

#### چکیده

هدف از این تحقیق ارزیابی کارایی سه الگوریتم ماشین بردار پشتیبان، درخت تصمیم‌گیری و شبکه عصبی مصنوعی فازی آرتمپ جهت تهیه نقشه پوشش زمین حوزه آبخیز ارکواز با استفاده از تصویر سنجندهOLI ماهواره لندست ۸ (سال ۱۳۹۳) است. پس از تصحیح هندسی تصاویر و پیش‌پردازش‌های صورت گرفته، اقدام به تعیین نمونه‌های تعلیمی مربوط به طبقات پوشش زمین برای انجام عملیات طبقه‌بندی شد و ارزیابی تفکیک پذیری نمونه‌ها در هر کدام از طبقات پوشش زمین با استفاده از شاخص آماری واگرایی تعیین شد. به منظور ارزیابی صحت نتایج طبقه‌بندی هر کدام از الگوریتم‌ها، نقشه واقعیت زمینی با استفاده از روش سیستماتیک به ابعاد ۵۵۰ متری طراحی و نوع پوشش زمینی آن‌ها در محل نمونه‌ها مشخص گردید. در نهایت، کارایی روش‌های طبقه‌بندی با استفاده از معیارهای صحت کلی، ضریب کاپا، صحت تولید کننده 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 نشان داد که طبقه‌بندی کننده ماشین بردار پشتیبان با مقدار صحت کلی ۰/۹۱/۲۶٪ و ضریب کاپای ۰/۸۷۳۱٪ نتایج مطلوب‌تری نسبت به سایر الگوریتم‌ها دارد. نتایج همچنین نشان داد که تفکیک و طبقه‌بندی اراضی جنگلی از سایر طبقات کاربری منطقه با صحت بالا صورت گرفته است.

\* Vahidmirzaei6764@gmail.com: پست الکترونیکی مسئول مکاتبات

## مقدمه

تهیه نقشه پوشش اراضی در مدیریت منابع طبیعی و محیط‌زیست، تهیه برنامه استفاده از سرزمین و تعیین توان و استعداد اراضی ضروری است و به عنوان یکی از منابع اصلی در تدوین برنامه‌های توسعه محسوب می‌شوند (۲). علم و آگاهی مدیران و کارشناسان از کم و کیف پوشش اراضی به جهت سیاست‌گذاری و چاره‌اندیشی برای رفع مشکلات موجود ضرورتی انکارناپذیر است (۱). تهیه نقشه پوشش اراضی، به خصوص در آن بخش از حوزه‌های آبخیز که دسترسی زمینی به آن‌ها به سختی میسر می‌شود، بسیار مشکل بوده و در بیشتر پروژه‌ها با محدودیت‌های مالی و زمانی همراه می‌باشد (۹). برای کشف و ارزیابی پوشش اراضی، سنجش از دور به عنوان علم و فن اطلاعات مکانی و سیستم اطلاعات جغرافیایی به خاطر برخورداری از امکانات تحلیلی می‌تواند نقش اساسی داشته باشند با استفاده از تصاویر سنجش از دور می‌توان داده‌های متنوع، ارزان‌تر، به هنگام، با کیفیت بهتر و البته با تکرار زیاد را برای اهداف مختلف به دست آورد (۱۹ و ۲۰). تجزیه و تحلیل این داده‌ها می‌تواند بینش‌های صحیح جهت تعامل انسان با محیط طبیعی فراهم کند به خصوص استفاده از تجزیه و تحلیل تصاویر چند طیفی می‌تواند به انسان جهت شناسایی پوشش زمین کمک کند (۳۲). از جمله کاربردهای داده‌های سنجش از دور تهیه نقشه‌های کاربری و پوششی است (۲۹). یکی از تحلیل‌های مهمی که بر روی تصاویر ماهواره‌ای انجام می‌گیرد، طبقه‌بندی این تصاویر می‌باشد. در حقیقت فرآیند طبقه‌بندی تصویر، تبدیل داده‌ها به اطلاعات قابل درک است (۲۴). طبقه‌بندی رقومی تصاویر کمک می‌کند تا فرآیند طبقه‌بندی به واقعیت نزدیک‌تر باشد (۲۵ و ۲۷). از آنجا که هدف اصلی از پردازش تصاویر ماهواره‌ای، تهیه نقشه‌های موضوعی و کارآمد می‌باشد، انتخاب الگوریتم مناسب طبقه‌بندی نقش زیادی در این امر ایفاء 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

روش‌های طبقه‌بندی پارامتریک (مانند حداکثر احتمال و حداقل فاصله)، وابستگی آن‌ها به توزیع آماری داده‌ها (توزیع نرمال گوسن) می‌باشد (۲۳). به همین دلیل روش‌های غیر پارامتریک، از جمله روش‌های مبتنی بر شبکه‌های عصبی مصنوعی و درخت تصمیم‌گیری و ماشین بردار پشتیبان برای طبقه‌بندی تصاویر ماهواره‌ای بکار گرفته شده است.

اگرچه روش‌های شبکه عصبی مصنوعی ممکن است با دقت بالا جهت طبقه‌بندی تصاویر ماهواره‌ای مورد استفاده قرار گیرد، اما تعدادی از مطالعات نشان داده‌اند که استفاده کنندگان از روش‌های شبکه عصبی 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 (۳۳ و ۳۹).

الگوریتم طبقه‌بندی دیگر، درخت تصمیم‌گیری 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 ساده‌تر تبدیل می‌نماید. روش‌های درخت تصمیم‌گیری مزیت‌هایی برای طبقه‌بندی تصاویر ماهواره‌ای بخاطر انعطاف‌پذیری، ماهیت غیر پارامتریکی و توانایی اجرا روابط غیر پارامتریکی بین عوارض و طبقه‌ها دارند. دقت یک طبقه‌بندی انجام شده به وسیله 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.

در چند سال اخیر یک تکنیک جدید بر اساس تئوری یادگیری آماری با نام ماشین‌های بردار پشتیبان (SVM) جهت طبقه‌بندی داده‌های سنجش از دور مورد استفاده واقع شده است (۱۴ و ۳۷). ماشین‌های بردار پشتیبان الگوریتم‌های سنجش نظارت شده بر اساس تئوری سنجش آماری هستند. هدف الگوریتم SVM برای طبقه‌بندی، تعیین فاصله‌های است که دو طبقه را به بهترین نحو جدا سازد. یک فاصله‌بینه با

احتمال دقت بالاتری دارد. ماشین بردار پشتیبان یکی دیگر از روش‌های طبقه‌بندی تصاویر ماهواره‌ای است که در این تحقیق مورد بررسی قرار گرفته است. روش ماشین‌های بردار پشتیبان یک روش طبقه‌بندی الگو است که اولین بار توسط وپنیک (۳۳) معرفی شد.

کونرن و همکاران (۱۷) برای طبقه‌بندی جنگل از روش ماشین بردار پشتیبان استفاده کردند. هدف آن‌ها از این تحقیق، ارزیابی دقت طبقه‌بندی در مجموعه از تصاویر متوالی بود. این تحقیق نشان داد که این روش زمانی که سطح زمین از چندین عارضه پوشیده شده باشد، نیز کارآمد خواهد بود.

کندوبی و همکاران (۱۸) در پژوهشی پیرامون تنوع و بیوماس، از تصاویر آیکنوس برای مدل‌سازی و پیش‌بینی این پدیده استفاده کردند. آن‌ها ۶ روش مختلف طبقه‌بندی را با همدیگر مقایسه کردند و به این نتیجه رسیدند که روش ماشین بردار پشتیبان بالاترین رتبه را میان سایر الگوریتم‌های بکار رفته دارا است.

وانگ و همکاران (۳۴) برای تهیه نقشه واحدهای لیتولوژیک و کاهش مشکل حضور پوشش گیاهی در تصویر، از ترکیب روش ماشین بردار پشتیبان با بر روی سنجنده‌های TM<sup>+</sup> و ETM<sup>+</sup> استفاده کردند و روش SVM-ACA را به عنوان روشی کارآمد در نواحی دارای پوشش گیاهی معرفی نمودند.

مونتراکیس و همکاران (۲۴) در پژوهشی به بررسی روش SVM پرداختند و تحقیقات انجام گرفته در این حوزه را در قالب هدف استفاده، آشکارسازی اهداف ایستا و پویا، قدرت تفکیک مکانی و طیفی تصاویر، نوع تصاویر و حیطه‌های کاربرد این روش، طبقه‌بندی کردند.

نیک نژاد و همکاران (۲۸) در مطالعه‌ای اقدام به مقایسه الگوریتم‌های مختلف طبقه‌بندی تصاویر ماهواره‌ای در تهیه نقشه جنگل در جنگل‌های بیوره استان ایلام نمودند که نتایج نشان داد که الگوریتم ماشین بردار پشتیبان نسبت به روش‌های دیگر با ضریب کاپای ۰/۷۰ و صحت کلی ۸۸/۶۵ درصد دارای دقت بالاتری است. تحقیقاتی که در داخل کشور صورت گرفته است با استفاده از روش‌های معمول طبقه‌بندی مثل الگوریتم

استفاده از سری نمونه‌های تعلیمی و توانایی تعمیم آن‌ها تعیین می‌شود. ایده اصلی در این روش طبقه‌بندی، تفکیک طبقه‌ها با یک سطحی است که مرز بین طبقه‌ها با استفاده از پیکسل‌های مرزی به حداکثر برسد. SVM به طور مستقیم از طریق یک فرآیند بهینه‌سازی، به جستجوی سطح جدا کننده (فراصفحه) بین طبقه‌های طیفی می‌پردازند. این سطح جداکننده یا مرز بین طبقه‌ها، با استفاده از نمونه‌های آموزشی محدودی که نزدیکترین فاصله را تا فراصفحه دارند و به بردارهای پشتیبان معروف هستند، مشخص می‌شود.

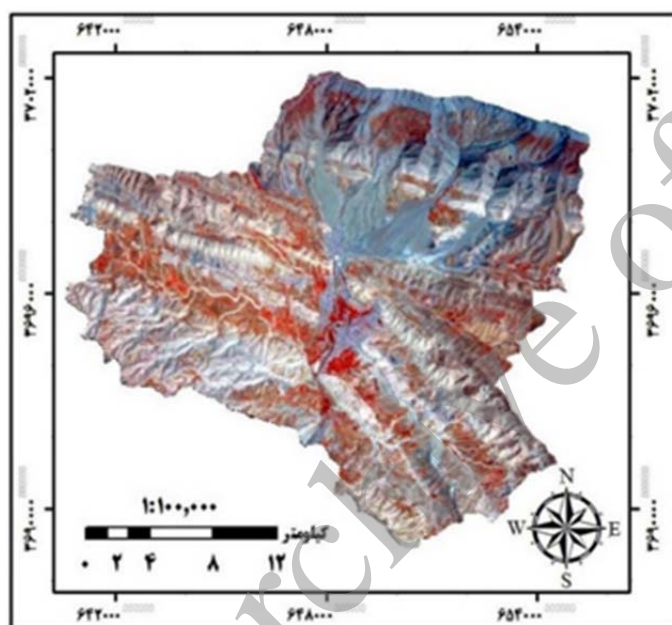
از جمله مطالعات انجام شده بر روی روش‌های طبقه‌بندی تهیه نقشه پوشش سطح زمین با قدرت تفکیک ۱ کیلومتر با استفاده از رویکرد طبقه‌بندی نظارت نشده توسط لولند و همکاران (۲۱) انجام گرفته است. به علت محدودیت‌های این روش‌ها، در سال‌های اخیر، رویکردهای طبقه‌بندی درختی و شبکه عصبی مصنوعی که خصوصیات غیرخطی و غیر پارامتریکی دارند، در سطح منطقه‌ای و جهانی مورد استفاده قرار گرفتند. هانسن و همکاران (۱۲) از داده‌های NOAA/AVHRR برای تهیه نقشه پوشش سطح زمین در مقیاس جهانی و روش‌های حداکثر احتمال و طبقه‌بندی درختی استفاده کرده‌اند. در این بررسی دقت روش طبقه‌بندی درختی نسبت به حداکثر احتمال بهتر گزارش شده است. شبکه عصبی روش دیگری می‌باشد که به طور گسترده در طبقه‌بندی داده‌های سنجنش از دور مورد بررسی واقع شده است. برای غلبه به مشکل ساختاری شبکه عصبی، روش آرتمپ فازی توسعه یافت که این روش می‌تواند منطق فازی و تئوری تشدید انطباقی را با هم تلفیق کند. گوپال و همکاران (۱۱) با استفاده از رویکرد آرتمپ فازی پوشش سطح زمین را در مقیاس جهانی طبقه‌بندی نموده‌اند. بوراک و همکاران (۳) روش‌های حداکثر احتمال، درخت تصمیم‌گیری و شبکه عصبی 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 تصمیم‌گیری و حداکثر

## مواد و روش‌ها

### منطقه مورد مطالعه

حوزه آبخیز ارکواز با مساحت ۱۲۰۷۹/۳ هکتار در استان ایلام، در شهرستان ملکشاهی قرار گرفته و از نظر موقعیت جغرافیایی بین  $30^{\circ} 46'$  تا  $33^{\circ} 23'$  طول شرقی و  $33^{\circ} 19'$  تا  $33^{\circ} 52'$  عرض شمالی واقع شده است. حداکثر ارتفاع حوزه از سطح دریا ۲۷۳۷/۳ متر و حداقل ارتفاع در خروجی حوزه برابر ۱۲۲۶/۰۳ متر از سطح دریا می‌باشد. شکل ۱ موقعیت حوزه را در کشور و استان ایلام به تصویر کشیده است.

حداکثر احتمال، حداقل فاصله از میانگین و متوازی السطوح بوده است (۱). تا به حال مطالعات اندکی با استفاده از روش‌های مورد اشاره در این مطالعه انجام شده است، لذا در این مطالعه از الگوریتم‌های ماشین‌های بردار پشتیبان، درختی و شبکه عصبی مصنوعی آرتمپ فازی برای طبقه‌بندی پوشش اراضی کوهستانی حوزه آبخیز ارکواز و با استفاده از تصویر ماهواره‌ای OLI کسب شده در سال ۱۳۹۳ استفاده شده است. همچنین کارآمدی این روش‌ها با هم مقایسه و بصورت آماری آنالیز گردیده است.



شکل ۱. موقعیت حوزه آبخیز ارکواز در ایران و استان ایلام

اساس تصویر تولیدی در این مرحله انجام می‌گیرد. نوع و نحوه انجام این عملیات بسته به عوامل مختلفی چون نوع داده‌های مورد استفاده و هدف تحقیق، متغیر خواهد بود. به منظور بررسی وجود و یا عدم وجود خطاهای هندسی و رادیومتری، بررسی کیفیت داده‌ها بر روی تصاویر ماهواره‌ای صورت گرفت. با بررسی تصاویر تک باند و ترکیبات رنگی مختلف هیچ گونه خطای قابل ملاحظه‌ای مشاهده نشد. برای تصحیح هندسی از نقاط کنترل زمینی به دست آمده از روی نقشه

### داده‌های مورد استفاده

در این تحقیق از داده‌های تصاویر سنجنده OLI ماهواره لندست ۸ مربوط به تاریخ ۱۳۹۳/۰۴/۳۱ و نقشه توپوگرافی با مقیاس ۱:۲۵۰۰۰ استفاده گردید.

### پیش‌پردازش تصاویر

مرحله پیش‌پردازش داده‌ها یکی از مهم‌ترین مراحل در پردازش تصاویر است، چرا که تمامی محاسبات بعدی بر

طبقه‌بندی نیز مهم‌تر باشد (۱۴). انتخاب نمونه‌های تعلیمی مورد نیاز برای طبقه‌بندی طبقه‌های مختلف در منطقه با شناخت از منطقه و با استفاده از اطلاعات جنبی و استفاده از تصویر رنگی مرکب ۲، ۳ و ۴ (RGB) صورت گرفت. انتخاب این تصویر رنگی مرکب صرفاً جهت شناسایی طبقه‌های مورد نظر و انتخاب بهترین نمونه‌های تعلیمی بوده است. برای این منظور در منطقه مورد مطالعه به فراخور سهم هر طبقه تعداد مناسبی نمونه تعلیمی به طور تصادفی با استفاده از بررسی‌های میدانی، تصاویر گوگل ارث 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 طبقه‌های فوق در باندهای اصلی و پردازش شده استخراج شد و نمودار توزیع ارزش طیفی نمونه‌های تعلیمی در تمامی باندهای طیفی بررسی شد. برای بررسی تشابه آنها و میزان تفکیک‌پذیری و تباین طبقه‌ها، از روش ارزیابی کمی تفکیک‌پذیری استفاده و تفکیک‌پذیری آنها با استفاده از شاخص واگرایی بررسی شد.

#### انتخاب بهترین مجموعه بانندی

با هدف انتخاب بهترین مجموعه باندهایی که تفکیک‌پذیری 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.

#### طبقه‌بندی تصاویر

پس از انتخاب بهترین ترکیبات بانندی عملیات طبقه‌بندی با استفاده از الگوریتم‌های ماشین بردار پشتیبان، شبکه عصبی مصنوعی آرتمپ فازی و روش درختی صورت گرفت. به منظور حذف پیکسل‌های منفرد و پراکنده در سطح تصویر طبقه‌بندی شده و همچنین به دست آوردن تصویر مطلوب و با وضوح بیشتر، از فیلتر پایین گذر نما استفاده شد.

توپوگرافی ۱:۲۵۰۰۰ و نقاط مرجع زمینی با استفاده از GPS استفاده شد. همچنین عمل نمونه‌گیری مجدد با استفاده از روش نزدیکترین همسایه به منظور جلوگیری از تغییر ارزش‌های طیفی تصاویر بکار گرفته شد. پس از بکارگیری روش ناپارامتری چندجمله‌ای و حذف نقاط نامناسب، تصحیح هندسی با نقاط باقیمانده و میزان خطای ریشه مربعات ۰/۲۱ پیکسل صورت گرفت. پس از آنکه داده‌ها مورد تصحیح هندسی قرار گرفتند، برای بهره‌مندی از توان اطلاعاتی داده‌ها، به کمک الگوریتم‌های طبقه‌بندی، قابلیت آن‌ها برای تفکیک طبقه‌های موضوعی که در تحقیق مورد نظر هستند، مورد بررسی و آزمون قرار گرفتند. بر اساس تجربه و نتایج تحقیقات گذشته، باندهای مصنوعی نظیر مؤلفه‌های اصلی و نسبت‌گیری‌هایی مثل NDVI و RVI به همراه باندهای اصلی سنجنده OLI می‌تواند برای تفکیک بهتر طبقه‌ها مورد استفاده قرار گیرند. ایجاد این باندها خود نوعی عملیات بارزسازی بوده که در نتیجه منجر به تغییر DNS شده و جزء عملیات قبل از طبقه‌بندی محسوب می‌شوند. شاخص‌های مزبور پوشش گیاهی با استفاده از تصاویر سنجنده OLI تولید شدند. در این تحقیق از روش تجزیه مؤلفه اصلی به روش استاندارد نیز استفاده شد. 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 بانندی برای عمل طبقه‌بندی مورد استفاده قرار گرفتند (۲۱).

#### طبقه‌بندی و تهیه نقشه پوشش اراضی

##### تهیه نمونه‌های تعلیمی

اساساً طبقه‌بندی تصاویر سنجش از دور فرآیند مشکلی است، چرا که بیشتر روش‌های طبقه‌بندی نظارت شده نیازمند داده‌های تعلیمی به اندازه کافی بزرگ هستند (۶). از طرف دیگر، نحوه انتخاب نمونه‌های تعلیمی نیز یکی از مهم‌ترین عوامل در طبقه‌بندی پیکسل‌ها است (۲۸). برای رسیدن به دقت بهتر، این عامل می‌تواند حتی از انتخاب الگوریتم

### الگوریتم‌های طبقه‌بندی کننده تصاویر

#### روش شبکه عصبی مصنوعی فازی آرتمپ

از روش‌های جدید در امر طبقه‌بندی تصاویر ماهواره‌ای است و به دو صورت نظارت شده و نظارت نشده می‌باشد. در شبکه نظارت شده، ورودی به سیستم اعمال و اختلاف بین خروجی شبکه و خروجی مطلوب برای تغییر و تطبیق وزن‌ها بکار می‌رود. در صورتی که در شبکه نظارت نشده، شبکه خود را متناسب با داده‌های ورودی تطبیق می‌کند. در این مطالعه، روش نظارت شده مورد استفاده قرار گرفت. شبکه عصبی مصنوعی آرتمپ فازی بر پایه نظریه تشدید انطباقی می‌باشد. ساختار شبکه‌های مبتنی بر نظریه تشدید انطباقی با آموزش با نظارت و به نام آرتمپ شناخته شده‌اند (۵).

#### روش طبقه‌بندی درخت تصمیم‌گیری

روش‌های طبقه‌بندی معمولی مانند بیشترین شباهت، حداقل فاصله و متوازی‌السطوح روش‌های یک مرحله‌ای نامیده می‌شوند و برای هر پیکسل فقط یک برچسب را تعیین می‌کنند. برخلاف این روش‌ها، روش‌های چند مرحله‌ای یا سلسله مراتبی قادر هستند که برچسب هر طبقه را در چندین مرحله تصمیم‌گیری تعیین کنند. در این روش، تصویر در سطوح مختلف طبقه‌بندی می‌شود. به همین دلیل خطای تجمعی مینیمم می‌شود که این موضوع یکی از مهم‌ترین مزایای این روش می‌باشد. طبقه‌بندی درختی مورد استفاده در این مطالعه، یک روش تک متغیره غیرپارامتریک می‌باشد. در عمل، این الگوریتم مشخصه‌ای (از قبیل باند انعکاسی) را به صورت 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 می‌گویند. سپس کل تصویر بر اساس همین داده‌های تعلیمی طبقه‌بندی می‌شود. برای شروع کار، تمام پیکسل‌های تعلیمی از تمام طبقه‌ها به ریشه اختصاص

می‌یابند. از آنجائیکه ریشه حاوی تمام پیکسل‌های تعلیمی از همه طبقه‌ها است، یک فرایند تکراری برای رشد درخت آغاز می‌شود طبقه‌ها را از یگدیگر جدا می‌سازد (۲۱).

در طبقه‌بندی درختی، به طور متوالی داده‌ها در هر میانگه به میانگه‌های جدید حاوی زیر مجموعه‌های همگن‌تر پیکسل‌های تعلیمی پیش می‌روند. یک میانگه جدید ایجاد شده ممکن است در حالتی که پیکسل‌های تعلیمی حاوی تنها یک طبقه است یا غالبیت پیکسل‌ها با یک طبقه است، ایجاد یک برگ نماید. زمانی که دیگر میانگه‌ای برای انشعاب (تفکیک) وجود ندارد، قواعد نهایی طبقه‌بندی درختی شکل می‌گیرد. مرحله بعدی طبقه‌بندی درختی، طبقه‌بندی تصویر است. در این مرحله، هر پیکسل با استفاده از قواعد تصمیم‌گیری از پیش تعریف شده به یک طبقه اختصاص می‌یابد. یک پیکسل ابتدا به ریشه درخت اختصاص می‌یابد. ارزش پیکسل با مقادیر ریشه سنجیده شده و به یک میانگه فرستاده می‌شود. فرآیند ادامه می‌یابد تا اینکه پیکسل به یک برگ می‌رسد و سپس برچسب طبقه را می‌گیرد. در این مطالعه، جهت هرس کردن درخت طبقه‌بندی شده از روش هرس خودکار استفاده شد. طبقه‌بندی درختی دارای الگوریتم‌های انشعاب آنتروپی، نسبت بهره و جینی می‌باشد. چون در مطالعات قبلی روش انشعاب جینی به عنوان بهترین روش معرفی شده بود (۱۲) در تحقیق حاضر از روش انشعاب جینی استفاده شده است. قانون انشعاب جینی معیار ناخالصی در یک میانگه معین بوده و مقدار آن در حالتی که همه پیکسل‌ها در تمام طبقه‌ها توزیع یکنواخت داشته باشند، در حداکثر خود می‌باشد. به طور کلی، قاعده انشعاب جینی سعی می‌کند همگن‌ترین دسته را در میان سری داده‌ها پیدا کند و آن را از بقیه داده‌ها جدا کند (۳۸).

#### الگوریتم ماشین بردار پشتیبان

این الگوریتم، بر پایه تئوری یادگیری آماری که توسط واپنیک، ابداع شد، بنا شده است (۳۳). بر اساس این تئوری می‌توان کران نرخ خطای ماشین یادگیری را برای داده‌های

بیشینه و مقدار خطا را کمینه کند. این فرا صفحه، با حل مسئله بهینه‌سازی قید دار زیر حل می‌شود.

$$\min \left\{ \frac{1}{2} \|w\|^2 + C \sum_{i=1}^k \xi_i \right\} \quad [1]$$

$$y_i (w \cdot x_i + b) \geq 1 - \xi_i \quad i = 1, \dots, k$$

در این رابطه؛ پارامتر ثابت  $C$ ، که در آن  $0 < C < \infty$ ، پارامتر تنظیم است، این پارامتر، تعادلی بین اثر نمونه‌های آموزشی بدون طبقه‌بندی و حاشیه بیشینه، برقرار می‌کند و مقدار آن با سعی و خطا بدست می‌آید. مسئله بهینه‌سازی قید دار بالا با روش ضرائب لاگرانژ، به صورت زیر قابل حل می‌باشد.

$$\text{Maximize: } \sum_i a_i - \frac{1}{2} \sum_{i,j} a_i a_j y_i y_j K(x_i \cdot x_j) \quad [2]$$

$$\text{Subject to: } \sum_{i=1}^n a_i y_i = 0 \text{ and } 0 \leq a_i \leq C \quad i = 1, \dots, k$$

در این رابطه،  $a_i \geq 0$ ، ضرایب لاگرانژ هستند و مسئله بهینه‌سازی فوق به محاسبه این ضرایب منجر می‌شود. بر اساس شرایط بهینه‌سازی (Karush Kuhn Yucker) KKY برخی از ضرایب لاگرانژ دارای مقدار صفر و ضرایب غیر صفر، متناظر با بردارهای پشتیبان هستند (۱۹). حل این مسئله، مجموع جواب  $a^0 = (a_1^0, \dots, a_k^0)$  است و مقادیر  $b$  و  $w$  نیز از روابط  $w^a = \sum_i y_i a_i^0 x_i$  و  $b^0 = \frac{1}{2} [w^0 \cdot x_{+1}^0 + w^0 \cdot x_{-1}^0]$  محاسبه می‌شوند که در آن  $x_{+1}^0$  و  $x_{-1}^0$  متناظر با بردارهای پشتیبان طبقه‌های با برچسب  $+1$  و  $-1$  می‌باشند. در آخر تابع تصمیم‌گیری، برای طبقه‌بندی از رابطه ۳ به دست می‌آید.

$$f(x) = \text{sign} \left( \sum_{sv} y_i a_i^0 (x_i \cdot x) + b^0 \right) \quad [3]$$

با این وجود مواردی وجود دارد که در آن با یک فراصفحه خطی، نمی‌توان داده‌ها را از یکدیگر جدا کرد. در این موارد داده‌ها به فضای بعد بالاتر منتقل می‌شوند که در این فضا، با یک فراصفحه خطی، قابل جداسازی باشند. برای این کار، یک تابع تبدیل غیر خطی  $\Phi$  به فرم یک تابع هسته  $k$  با

طبقه‌بندی نشده، به عنوان نرخ خطای تعمیم یافته، در نظر گرفت. این کران‌ها به عنوان تابعی از مجموع نرخ خطاهای آموزشی هستند که میزان پیچیدگی طبقه‌بندی کننده‌ها را نشان می‌دهند. به منظور کمینه کردن نرخ خطاهای تعمیم یافته، بایستی نرخ خطای آموزش و پیچیدگی طبقه‌بندی کننده کاهش یابد. این کار را می‌توان با بیشینه کردن حاشیه جداسازی انجام داد. الگوریتم ماشین بردار پشتیبان دارای ۴ ویژگی اساسی است؛ (۱) قابلیت تعمیم بالا، برای جداسازی نمونه‌هایی که به دو طبقه مختلف تعلق دارند، از یک فرا صفحه با حاشیه بیشینه استفاده می‌کند که به آن فراصفحه تصمیم‌گیری بهینه گفته می‌شود. (۲) متغیر کمکی، هنگامی که داده‌ها به صورت خطی قابل جداسازی نیستند، SVM این قید را با استفاده از متغیر کمکی خطای مجاز برای هر نمونه در نظر می‌گیرد. (۳) فضای ویژگی، SVM با تبدیل داده‌ها از فضای  $R^N$  به فضای ویژگی با ابعاد  $R^H$  که در آن  $H > N$ ، یک تابع تصمیم‌گیری غیر خطی ایجاد می‌کند که در فضای ویژگی حاصل از آن، داده‌ها به صورت خطی قابل جدا شدن هستند. (۴) بردارهای پشتیبان، فقط بردارهای پشتیبان هستند که در ساخت فرا صفحه نقش دارند. در یک مسئله طبقه‌بندی دوتایی، برای یک مجموعه داده که به دو طبقه تعلق دارند به دنبال یافتن یک فرا صفحه خطی هستیم. فرض کنیم که مجموعه آموزشی، شامل  $K$  نمونه آموزشی  $\{(x_1, y_1), \dots, (x_K, y_K)\}$  است و  $x_i \in R^N$  و هر نمونه به یکی از دو طبقه  $\{+1, -1\}$  تعلق دارد. هدف روش SVM، پیدا کردن یک تابع تصمیم‌گیری خطی با رابطه  $f(x) = w \cdot x + b$  می‌باشد که در آن  $w \in R^N$  جهت فرا صفحه را مشخص می‌کند و  $b \in R$  پارامتر تمایل به یک طرف فرا صفحه می‌باشد. معادله فرا صفحه برای دو طبقه با رابطه  $y_i(w \cdot x_i + b) \geq 1$  مشخص می‌شود. به علت وجود خطا و ترکیب طبقه‌ها در هنگام انتخاب نمونه‌های آموزشی، از متغیر آزاد  $\xi_i$  جهت در نظر گرفتن اثر نمونه‌های آموزشی بدون طبقه‌بندی، استفاده می‌شود. بنابراین، معادله فرا صفحه برای دو طبقه، در این حالت با رابطه  $y_i(w \cdot x_i + b) \geq 1 - \xi_i$  مشخص می‌شود. فرا صفحه بهینه، در مکانی واقع است که حاشیه بین دو طبقه را

رابطه زیر در نظر گرفته می‌شود.

$$k(X_i, X_j) = \phi(X_i)\phi(X_j) \quad [4]$$

در این رابطه، جایگزین ضرب نقطه‌ای در فضای جدید، بین بردارها می‌شود. مسئله بهینه‌سازی، در این حالت به صورت زیر تغییر می‌یابد.

$$\text{Maximize} : \sum_i a_i - \frac{1}{2} \sum_{i,j} a_i a_j y_i y_j K(X_i, X_j) \quad [5]$$

$$\text{Subject to} : \sum_{i=1}^n a_i y_i = 0 \text{ and } 0 \leq a_i \leq C \quad i = 1, \dots, k$$

که تابع تصمیم‌گیری نیز از رابطه ۶ به دست می‌آید.

$$f(x) = \text{sign} \left( \sum_{SV} y_i a_i^0 K(x_i, x) + b^0 \right) \quad [6]$$

هسته‌های مورد استفاده در این تابع، برای تبدیل داده‌ها از فضای ورودی با بعد بالاتر، معمولاً یکی از سه هسته؛ خطی  $(K(X_i, X_j) = X_i \cdot X_j)$ ، چند جمله‌ای  $(K(X_i, X_j) = (X_i \cdot X_j + 1)^d)$  و شعاعی  $(K(X_i, X_j) = \exp(-\gamma \|X_i - X_j\|^2))$  است (۴).

در این مطالعه از هسته شعاعی یا کرنل شعاعی به دلیل کارایی بیشتر استفاده شده است (۳۰). روابطی که در بالا تشریح شد برای یک طبقه‌بندی دو طبقه کارایی دارد، چرا که SVM، در فرم اولیه یک طبقه‌بندی کننده دوتایی است. با بکارگیری چند راهکار در الگوریتم اصلی SVM، می‌توان طبقه‌بندی‌های K طبقه را نیز با این روش پیاده‌سازی نمود. یکی از این راهکارها، ترکیب طبقه‌بندی کننده یکی در برابر بقیه می‌باشد، این روش آسان‌ترین روش است و نتایج منطقی و قابل قبولی را ارائه می‌دهد. در این روش برای هر طبقه، یک طبقه‌بندی کننده و در مجموع، K طبقه‌بندی کننده ساخته می‌شود. جهت آموزش K امین طبقه‌بندی کننده از نمونه‌های آموزشی طبقه K ام در مقابل سایر نمونه‌های آموزشی استفاده می‌شود. یکی دیگر از روش‌های طبقه‌بندی چند طبقه، ترکیب طبقه‌بندی کننده‌های یکی در برابر یکی می‌باشد، این روش

برای هر زوج طبقه ممکن نیاز به یک طبقه‌بندی کننده دوتایی دارد. تعداد کل طبقه‌بندی کننده‌ها برای K طبقه، برابر با  $K(K-1)/2$  می‌باشد. نمونه‌های آموزشی برای هر یک از طبقه‌بندی کننده‌ها، زیرمجموعه‌ای از نمونه‌های آموزشی موجود است که فقط شامل نمونه‌های آموزشی مربوط به آن دو طبقه بوده و به صورت +1 و -1 در یکی از دو طبقه برچسب‌گذاری خواهد شد. در نهایت این طبقه‌بندی کننده‌ها، با یکدیگر ترکیب شده تا نتایج نهایی طبقه‌بندی به دست آید (۱۳). در این مطالعه نیز از روش اول برای طبقه‌بندی طبقه‌های مختلف استفاده شد (به دلیل اینکه آسان‌ترین روش بوده و نتایج منطقی و قابل قبولی را ارائه می‌دهد) (۱۵).

#### تهیه نقشه واقعیت زمینی

در این تحقیق برای تعیین صحت نقشه‌های حاصل از الگوریتم‌های مختلف طبقه‌بندی داده‌های ماهواره‌ای، با استفاده از روش نمونه‌برداری تصادفی سیستماتیک، یک شبکه نمونه‌برداری ۵۵۰ 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. با توجه به اطلاعات به دست آمده از عملیات صحرائی و با بهره‌گیری از قابلیت سامانه‌های اطلاعات جغرافیایی نقشه واقعیت زمینی با پنج طبقه تهیه گردید. به دلیل ساختار رستری تصاویر ماهواره‌ای و نیز با توجه به اینکه این نقشه مبنای ارزیابی دقت طبقه‌بندی تصاویر مورد استفاده قرار می‌گیرد، ساختار نقشه واقعیت زمینی از حالت برداری به رستری تبدیل شد. در مجموع ۱۹۵ پلی‌گون از منطقه مورد مطالعه برداشت و از آنها یک نقشه با ساختار رستری تهیه شد. در نهایت نقشه‌های حاصل از الگوریتم‌های مختلف طبقه‌بندی با نقشه واقعیت زمینی مقایسه شده و ماتریس خطا، تشکیل شد.



## روش های ارزیابی الگوریتم های طبقه بندی

معیارهای بررسی عملکرد الگوریتم های طبقه بندی یا ارزیابی آنها شامل دقت، سرعت، پایداری و جامع بودن می باشد. اینکه کدام معیار یا کدام گروه از معیارها استفاده شوند، بستگی به هدف مطالعه دارد. معیاری که به تمام اهداف مربوط می باشد دقت یا صحت است که در این بررسی به عنوان معیار ارزیابی الگوریتم های مختلف انتخاب شد. سرعت و پایداری همچنین فاکتورهای مهم در ارزیابی کارایی الگوریتم ها هستند که در این بررسی مورد توجه قرار گرفتند. دو نوع از معیارهای پر استفاده تعیین دقت، یعنی صحت کلی و ضریب کاپا 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 معنی داری بین دو الگوریتم را فراهم می کند (۸).

## نتایج

تصحیح هندسی تصاویر با میزان خطای ریشه میانگین مربعات (RMSe) حدود ۰/۲۱ پیکسل ثبت شد و با رویهم گذاری لایه های خطی رودخانه ها بر روی تصویر تطابق یافته، دقت بالای عمل تطابق هندسی مورد تأیید قرار گرفت.

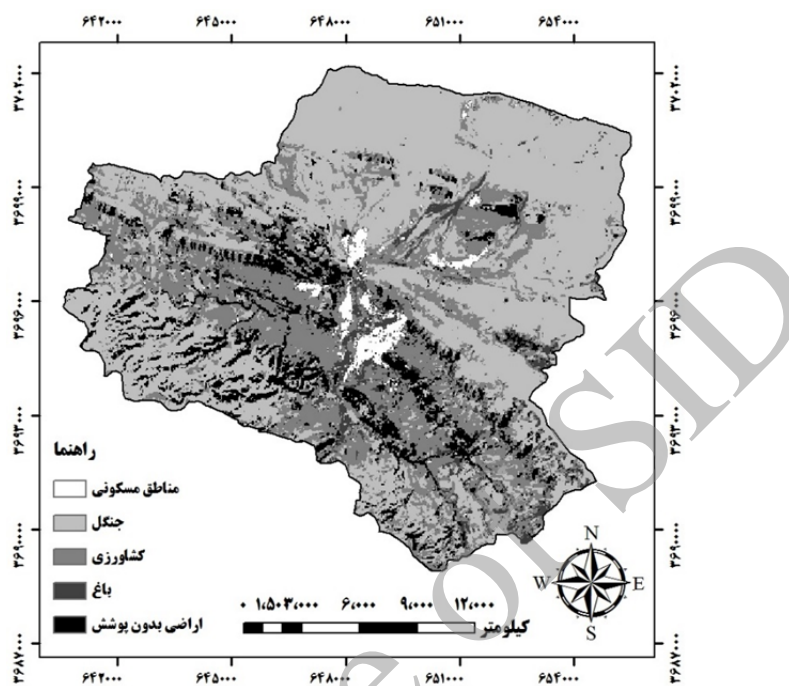
پس از تصحیح هندسی تصاویر و اطمینان از دقت زمین مرجع بودن تصاویر، از باندهای مصنوعی نظیر مؤلفه های اصلی و نسبت گیری های رایج (NDVI و RVI)، به همراه باندهای اصلی سنجنده OLI می تواند برای تفکیک بهتر طبقه ها مورد استفاده قرار گیرند. از بین مؤلفه های ایجاد شده ناشی از تجزیه مؤلفه های اصلی، مؤلفه های اول و دوم، به عنوان مناسب ترین مؤلفه ها انتخاب و در طبقه بندی شرکت داده شدند. با وجود ترکیبات مختلف باندهای حاصل از آنالیزهای مختلف، نتایج نشان دهنده مناسب بودن باندهای اصلی سنجنده OLI، برای طبقه بندی کاربری می باشد. جهت طبقه بندی تصویر ماهواره ای OLI، طبقه های پوشش زمینی در پنج گروه تحت عنوان طبقه های اراضی کشاورزی، باغات، جنگل، اراضی بایر و مناطق مسکونی تعیین و سپس نمونه های تعلیمی به صورت تصادفی از سطح منطقه با استفاده از تصاویر ماهواره ای گوگل ارث و بازدید میدانی جمع آوری شد. در مرحله بعد با استفاده از ویژگی های تصاویر، طبقه های پوشش اراضی در محدوده مورد مطالعه وارد شد و تفکیک پذیری طبقه ها با معیار فاصله واگرایی تبدیل شده و با استفاده از خصوصیات آماری نمونه های تعلیمی مورد بررسی قرار گرفت (جدول ۱).

جدول ۱. میزان تفکیک پذیری طبقات کاربری اراضی با استفاده از شاخص واگرایی

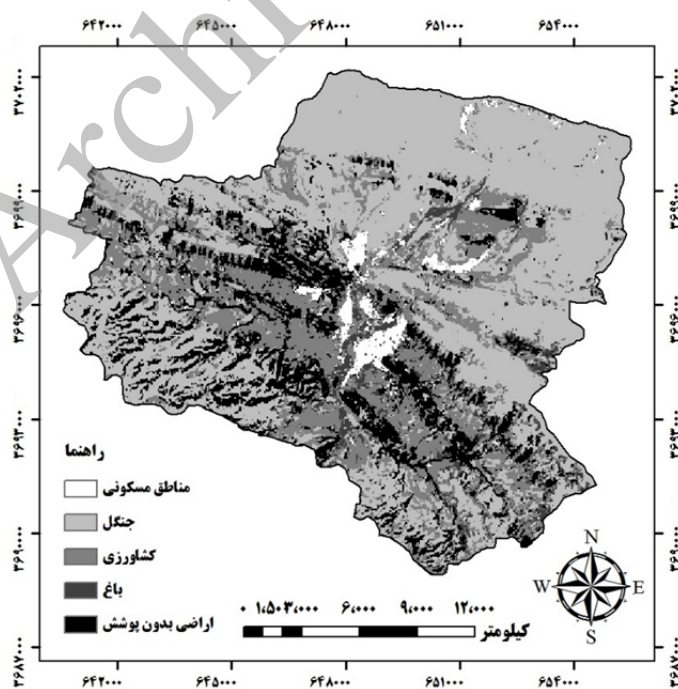
وضعیت تفکیک پذیری	میزان تفکیک پذیری	طبقات کاربری
خوب	۲/۰۰۰۰۰	مناطق مسکونی با اراضی کشاورزی
خوب	۲/۰۰۰۰۰	مناطق مسکونی با جنگل
کم	۱/۷۶۱۵۱	مناطق مسکونی با اراضی لخت
خوب	۱/۹۸۸۹۴	مناطق مسکونی با باغ
خوب	۱/۹۷۴۵۶	اراضی کشاورزی با جنگل
خوب	۲/۰۰۰۰۰	اراضی کشاورزی با اراضی لخت
خوب	۲/۰۰۰۰۰	اراضی کشاورزی با باغ
خوب	۲/۰۰۰۰۰	جنگل با اراضی لخت
خوب	۱/۹۹۸۵۲	جنگل با باغ
کم	۱/۷۷۴۶۸	اراضی لخت با باغ
خوب	۱/۹۴۹۸۲	میانگین تفکیک پذیری تمام طبقات

مصنوعی آرتف فازی (FA) و روش درختی (DT) اقدام شد. بدین ترتیب نقشه‌های پوشش اراضی مربوط به سال ۱۳۹۳ به دست آمد (شکل‌های ۲، ۳ و ۴).

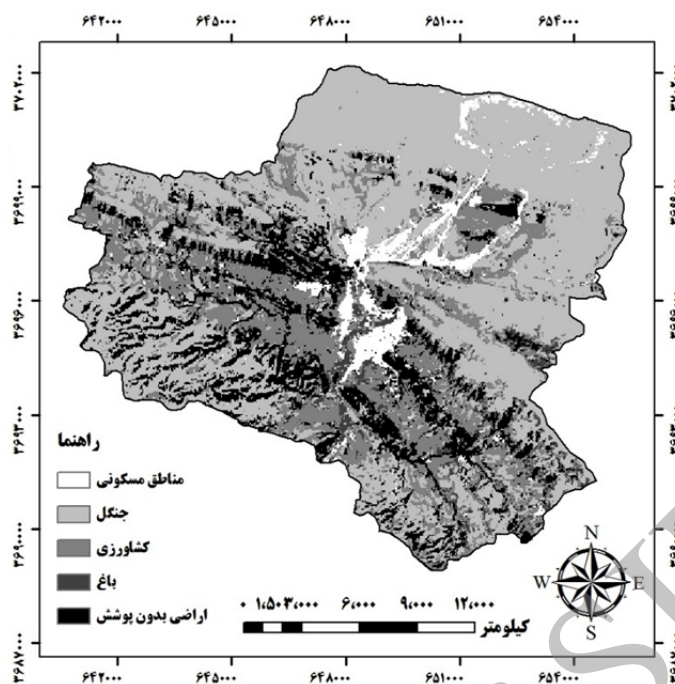
پس از مشخص نمودن میزان تفکیک‌پذیری طبقه‌ها، نسبت به طبقه‌بندی تصاویر سنجنده OLI در بهترین ترکیب بانندی به سه روش ماشین بردار پشتیبان (SVM)، شبکه عصبی



شکل ۲. تصویر طبقه‌بندی شده با استفاده از الگوریتم ماشین بردار پشتیبان



شکل ۳. تصویر طبقه‌بندی شده با استفاده از الگوریتم درخت تصمیم‌گیری



شکل ۴. تصویر طبقه‌بندی شده با استفاده از الگوریتم فازی آرتمپ

بردار پشتیبان (SVM) نسبت به سایر الگوریتم‌ها، پایدارتر می‌باشد. از دو الگوریتم دیگر، روش درختی (DT) نسبت به روش شبکه عصبی مصنوعی آرتمپ فازی (FA) بصورت محسوسی پایدارتر بوده است. همچنین از نظر سرعت اجراء الگوریتم روش درختی نسبت به دو الگوریتم دیگر سریع‌تر بوده و روش شبکه عصبی مصنوعی آرتمپ فازی کمترین سرعت را در اجرای عملیات طبقه‌بندی نشان داد.

نتایج حاصل از ارزیابی معیارهای صحت (تولید کننده و کاربر) در جدول ۲ و همچنین نتایج حاصل از طبقه‌بندی با استفاده از سه الگوریتم ماشین بردار پشتیبان، شبکه عصبی مصنوعی فازی آرتمپ و روش درختی با استفاده از مجموعه باندی مناسب در جدول ۳ ارائه شده است. به طور 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 ماشین

جدول ۲. نتایج ارزیابی صحت روش‌های مختلف به تفکیک طبقه‌ها

فازی آرتمپ		درخت تصمیم‌گیری		ماشین بردار پشتیبان		طبقات کاربری
صحت کاربر	صحت تولیدکننده	صحت کاربر	صحت تولیدکننده	صحت کاربر	صحت تولیدکننده	
۸۶/۴۱	۹۱/۹۳	۹۱/۵۰	۹۵/۰۹	۹۱/۱۱	۹۶/۲۴	جنگل
۷۶/۶۲	۸۹	۸۹/۶۰	۹۰/۱۵	۸۹/۸۳	۹۴/۳۴	اراضی کشاورزی
۸۴/۶۵	۸۳/۱۶	۲۳/۵۴	۸۵/۳۰	۹۳/۱۲	۹۶/۱۸	باغ
۸۸/۷۸	۸۶/۱۳	۹۳/۱۸	۹۲/۴۶	۹۳/۵۱	۹۵/۱۱	مناطق مسکونی
۷۳/۱۵	۸۳/۷۴	۸۶/۵۴	۹۶/۱۶	۹۵/۷۰	۹۵/۳۰	اراضی لخت

جدول ۳. نتایج کلی ارزیابی صحت الگوریتم‌های طبقه‌بندی

بهرترین ترکیب باندی	الگوریتم طبقه‌بندی کننده	صحت کلی (%)	ضریب کاپا
	ماشین بردار پشتیبان	۹۱/۲۶	۰/۸۷۳۱
باندهای اصلی سنجنده OLI	درخت تصمیم‌گیری	۸۷/۳۲	۰/۸۴۱۳
	آرتمپ فازی	۸۲/۱۳	۰/۸۰۱۱

### بحث و نتیجه‌گیری

در این مطالعه، پس از تصحیحات لازم و پیش پردازش‌های اولیه تصاویر، اقدام به طبقه‌بندی 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 سه روش طبقه‌بندی ماشین بردار پشتیبان، طبقه‌بندی درختی و شبکه عصبی مصنوعی فازی آرتمپ جهت طبقه‌بندی کاربری اراضی و پوشش گیاهی با استفاده از سری داده چند طیفی بود. نتایج بررسی تفکیک‌پذیری طبقات و انتخاب مجموعه باندی مناسب برای طبقه‌بندی تصویر و تهیه نقشه پوشش زمینی منطقه نشان داد که با وجود استفاده از مجموعه باندهای اصلی و باندهای پردازش شده حاصل از آنالیزهای مختلف (باندهای حاصل از تجزیه و تحلیل مؤلفه اصلی و نسبت‌گیری)، باندهای اصلی سنجنده OLI مناسب‌ترین مجموعه باندی برای تفکیک طبقه‌های مورد نظر بوده‌اند که نتیجه حاصله با نتایج شتایی و عبدی (۲) مطابقت دارد. نتایج بررسی تفکیک‌پذیری طبقات با معیار فاصله واگرایی تبدیل شده نشان داد که میزان تفکیک‌پذیری طبقات به جزء طبقه اراضی بدون پوشش با باغ 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. روش طبقه‌بندی درختی برای برطرف‌سازی بسیاری از مشکلات طبقه‌بندی مورد استفاده قرار گرفته است. اگرچه گزارش شده است که این روش کم و بیش نتایج طبقه‌بندی دقیق‌تری نسبت به روش‌های معمول ارائه می‌دهد،

اما انتخاب بهترین روش انشعاب مبحث مهمی است که کارایی این الگوریتم را تحت تأثیر قرار می‌دهد. در این مطالعه، از روش انشعاب جینی برای طبقه‌بندی طبقه‌های 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 (۳۶)، ژو و همکاران (۳۵)، چویی و همکاران (۷) در تحقیقات خود این الگوریتم طبقه‌بندی را دقیق‌ترین روش برای طبقه‌بندی تصاویر ماهواره‌ای معرفی کردند. الگوریتم ماشین بردار پشتیبان که اخیراً برای طبقه‌بندی تصاویر ماهواره‌ای مورد استفاده قرار گرفته است 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 به عنوان بهترین تابع کرنل برای طبقه‌بندی طبقه‌های اصلی کاربری زمین در منطقه استفاده شد. مقادیر کاپا و صحت کلی برای هر روش طبقه‌بندی مورد استفاده در این تحقیق با داده‌های سنجنده OLI محاسبه شد. این مطالعه نشان می‌دهد که الگوریتم ماشین‌های بردار پشتیبان با صحت کلی ۹۱/۲۶ درصد و ضریب کاپا ۰/۸۷۳۱ دارای کارایی بیشتر و دقت بالاتری نسبت به دیگر روش‌ها می‌باشد. روش طبقه‌بندی درختی با مقدار صحت کلی ۸۷/۳۲ درصد و ضریب کاپا ۰/۸۴۱۳ و شبکه عصب مصنوعی آرتمپ فازی با مقدار صحت کلی ۸۲/۱۳ درصد و ضریب کاپا ۰/۸۰۱۱ در درجه اهمیت بعدی قرار دارند. این نتایج به طور کلی تصدیق کننده مطالعات پیشین می‌باشد که در آن‌ها الگوریتم ماشین بردار پشتیبان دقیق‌تر از روش‌های شبکه عصبی مصنوعی و درخت تصمیم‌گیری است. شفری و رامل (۳۰)، فودی و ماتور

این تحقیق نشان داد که استفاده از روش‌های رایج طبقه‌بندی تصاویر ماهواره‌ای به منظور استخراج نقشه پوشش جنگلی مانند حداکثر احتمال و حداقل فاصله از میانگین که به طور عمده در تحقیقات و پروژه‌های اجرایی استفاده می‌شوند نسبت به روش ماشین بردار دارای دقت پایین‌تری می‌باشد.

### منابع مورد استفاده

1. سلمان ماهینی، ع.، آ. نادعلی، ج. فقهی و ب. ریاضی. ۱۳۹۱. طبقه‌بندی مناطق جنگلی استان گلستان به روش حداکثر احتمال با استفاده از تصاویر ماهواره‌ای ETM<sup>+</sup> سال ۲۰۰۱. مجله علوم و تکنولوژی محیط‌زیست، ۱۴(۳): ۴۷-۵۶.
2. شتابی، ش. و ا. عبدی. ۱۳۸۶. تهیه نقشه کاربری اراضی در مناطق کوهستانی زاگرس با استفاده از داده‌های سنجنده ETM<sup>+</sup> (منطقه مورد مطالعه: حوزه سرخاب خرم‌آباد لرستان). مجله علوم کشاورزی و منابع طبیعی، ۱۴(۱): ۱۳-۲۳.
3. Borak JS. 1999. Feature selection and land cover classification of a MODIS-like data set for a semiarid environment. *International Journal of Remote Sensing*, 20(5): 919-938.
4. Camps-Valls G, Gómez-Chova L, Calpe-Maravilla J, Martín-Guerrero JD, Soria-Olivas E, Alonso-Chordá L, Moreno J. 2004. Robust support vector method for hyperspectral data classification and knowledge discovery. *Geoscience and Remote Sensing, IEEE Transactions on*, 42(7): 1530-1542.
5. Carpenter GA, Grossberg S, Reynolds JH. 1991. ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network. *Neural networks*, 4(5): 565-588.
6. Chi M, Feng R, Bruzzone L. 2008. Classification of hyperspectral remote-sensing data with primal SVM for small-sized training dataset problem. *Advances in space research*, 41(11): 1793-1799.
7. Chubey MS, Franklin SE, Wulder MA. 2006. Object-based analysis of Ikonos-2 imagery for extraction of forest inventory parameters. *Photogrammetric Engineering & Remote Sensing*, 72(4): 383-394.
8. Congalton RG. 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37(1): 35-46.

(۱۰) و هوانگ و همکاران (۱۶) در تحقیقات خود به این نتیجه رسیدند. تحقیقات اخیر نشان داده است که ماشین بردار پشتیبان نسبت به روش‌های دیگر طبقه‌بندی دارای دقت بیشتری است. ویژگی اصلی این روش توانایی بالا در استفاده از نمونه‌های تعلیمی کمتر و رسیدن به دقت بالاتر در مقایسه با سایر روش‌های طبقه‌بندی است. (۲۲ و ۲۴). علاوه بر این دقت بالای الگوریتم ماشین بردار پشتیبان نسبت به دیگر روش‌ها را می‌توان این طور بیان کرد که این روش از لحاظ محاسباتی سریع بوده و همچنین این روش از قوانین بهینه‌سازی جهت مکان‌یابی مرزهای بهینه بین طبقه‌ها کاربری استفاده می‌کند در نتیجه جایگزین مناسبی برای سایر الگوریتم‌های معمول طبقه‌بندی می‌باشند. از مزایای دیگر روش ماشین بردار پشتیبان می‌توان به قابلیت تهیه نقشه طبقه‌بندی شده با دقت بالاتر با استفاده از نمونه‌های تعلیمی کوچک نسبت به روش‌های دیگر طبقه‌بندی اشاره کرد. یافته‌های تحقیق حاضر با نتایج کندوبی و همکاران (۱۸) که در پژوهشی پیرامون تنوع و بیوماس، از تصاویر آیکونوس برای مدل‌سازی و پیش‌بینی این پدیده استفاده کردند آن‌ها ۶ روش مختلف طبقه‌بندی را با هم دیگر مقایسه کردند و به این نتیجه رسیدند که روش ماشین بردار پشتیبان بالاترین رتبه را میان سایر الگوریتم‌های بکار رفته دارا است. سو و همکاران (۳۱) که در مطالعه‌ای با استفاده از داده‌های ماهواره‌ای MISR نشان دادند که با استفاده از روش‌های ماشین بردار پشتیبان و روش‌های حداکثر احتمال می‌توان حتی با دو سوم مجموعه داده‌ها به دقتی رسید که با کل داده‌ها قابل دستیابی است، مطابقت دارد. علاوه بر این نتایج تحقیق حاضر به طور کلی تصدیق کننده مطالعات پیشین می‌باشد که در آنها الگوریتم ماشین بردار پشتیبان دقیق‌تر از روش‌های دیگر طبقه‌بندی تصاویر ماهواره‌ای است. (۳۰) و فودی و ماتور (۱۰) در تحقیقات خود به این نتیجه رسیدند. از نتایج تحقیق حاضر می‌توان جهت تهیه نقشه پوشش جنگل با دقت بالاتر با استفاده از روش ماشین بردار پشتیبان در کارهای جنگلداری و ارزیابی زیست‌محیطی و منابع طبیعی در مناطق با شرایط مشابه استفاده نمود. همچنین نتایج

9. Dontree S. 2003. Land use dynamics from multi temporal remotely-sensed data: a case study Northern Thailand, proceedings of Map Asia conference, 13-15 October, Putra World Trade Center, Kuala Lumpur, Malaysia.
10. Foody GM, Mathur A. 2004. A relative evaluation of multiclass image classification by support vector machines. *Geoscience and Remote Sensing, IEEE Transactions on*, 42(6): 1335-1343.
11. Gopal S, Woodcock CE, Strahler AH. 1999. Fuzzy neural network classification of global land cover from a 1 AVHRR data set. *Remote sensing of Environment*, 67(2): 230-243.
12. Hansen M, Dubayah R, DeFries R. 1996. Classification trees: an alternative to traditional land cover classifiers. *International journal of remote sensing*, 17(5): 1075-1081.
13. Hastie T, Tibshirani R. 1998. Classification by pairwise coupling. *The Annals of Statistics*, 26(2): 451-471.
14. Hixon R, Ramirez MS, Villoch M. 1980. Aspects of morphometrics and reproduction of the squid *Ommastrephes pteropus* Steenstrup 1885 in the western Gulf of Mexico. *American Malacological Bulletin*, 46: 54-60.
15. Hong J-H, Min J-K, Cho U-K, Cho S-B. 2008. Fingerprint classification using one-vs-all support vector machines dynamically ordered with naïve Bayes classifiers. *Pattern Recognition*, 41(2): 662-671.
16. Huang C, Davis L, Townshend J. 2002. An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing*, 23(4): 725-749.
17. Knorn J, Rabe A, Radeloff VC, Kuemmerle T, Kozak J, Hostert P. 2009. Land cover mapping of large areas using chain classification of neighboring Landsat satellite images. *Remote Sensing of Environment*, 113(5): 957-964.
18. Knudby A, LeDrew E, Brenning A. 2010. Predictive mapping of reef fish species richness, diversity and biomass in Zanzibar using IKONOS imagery and machine-learning techniques. *Remote Sensing of Environment*, 114(6): 1230-1241.
19. Langford M, Bell W. 1997. Land cover mapping in a tropical hillsides environment: a case study in the Cauca region of Colombia. *International Journal of Remote Sensing*, 18(6): 1289-1306.
20. Lillesand TM, Kiefer RW, Chipman JW. 2004. *Remote Sensing and Image Interpretation*. John Wiley and Sons. New York, 784.
21. Loveland T, Reed B, Brown J, Ohlen D, Zhu Z, Yang L, Merchant J. 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing*, 21(6-7): 1303-1330.
22. Mantero P, Moser G, Serpico SB. 2005. Partially supervised classification of remote sensing images through SVM-based probability density estimation. *Geoscience and Remote Sensing, IEEE Transactions on*, 43(3): 559-570.
23. Mather P, Tso B. 2009. *Classification methods for remotely sensed data*. 2nd ED. Chapter 2-3, CRC Press. 376 pp.
24. Mountrakis G, Im J, Ogole C. 2011. Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3): 247-259.
25. Munyati C. 2004. Use of principal component analysis (PCA) of remote sensing images in wetland change detection on the Kafue Flats, Zambia. *Geocarto International*, 19(3): 11-22.
26. Niknejad M, Mirzaei zadeh V, Heydari M. 2014. Comparing different classifications of satellite imagery in forest mapping (Case study: Zagros forests in Iran). *International Research Journal of Applied and Basic Sciences*, 8(7): 1407-1415.
27. Oommen T, Misra D, Twarakavi NK, Prakash A, Sahoo B, Bandopadhyay S. 2008. An objective analysis of support vector machine based classification for remote sensing. *Mathematical Geosciences*, 40(4): 409-424.
28. Paola J, Schowengerdt R. 1995. A review and analysis of backpropagation neural networks for classification of remotely-sensed multi-spectral imagery. *International Journal of remote sensing*, 16(16): 3033-3058.
29. Richards JA, Jia X. 2006. *Remote Sensing Digital Image Analysis an Introduction*, 4th Edition, Chapter 1, Springer, Germany, Berlin, Heidelberg.
30. Shafri H, Ramle F. 2009. A Comparison of Support Vector Machine and Decision Tree Classifications Using Satellite Data of Langkawi Island. *Information Technology Journal*, 8(1): 64-70.
31. Su L, Huang Y, Chopping M, Rango A, Martonchik J. 2009. An empirical study on the utility of BRDF model parameters and

- topographic parameters for mapping vegetation in a semi-arid region with MISR imagery. *International Journal of Remote Sensing*, 30(13): 3463-3483.
32. Szuster BW, Chen Q, Borger M. 2011. A comparison of classification techniques to support land cover and land use analysis in tropical coastal zones. *Applied Geography*, 31(2): 525-532.
33. Vapnik V. 1995. *The Nature of Statistical Learning Theory*. Springer Science & Business Media. 314 pp.
34. Wang X, Niu R, Wu K. 2011. Lithology intelligent identification using support vector machine and adaptive cellular automata in multispectral remote sensing image. *Optical Engineering*, 50(7): 076201-076212.
35. Xu M, Watanachaturaporn P, Varshney PK, Arora MK. 2005. Decision tree regression for soft classification of remote sensing data. *Remote Sensing of Environment*, 97(3): 322-336.
36. Yang C-C, Prasher SO, Enright P, Madramootoo C, Burgess M, Goel PK, Callum I. 2003. Application of decision tree technology for image classification using remote sensing data. *Agricultural Systems*, 76(3): 1101-1117.
37. Yao X, Tham L, Dai F. 2008. Landslide susceptibility mapping based on support vector machine: a case study on natural slopes of Hong Kong, China. *Geomorphology*, 101(4): 572-582.
38. Zambon M, Lawrence R, Bunn A, Powell S. 2006. Effect of alternative splitting rules on image processing using classification tree analysis. *Photogrammetric Engineering & Remote Sensing*, 72(1): 25-30.
39. Zhang Z, Verbeke L, De Clercq E, Ou X, De Wulf R. 2007. Vegetation change detection using artificial neural networks with ancillary data in Xishuangbanna, Yunnan Province, China. *Chinese Science Bulletin*, 52(2): 232-243.

Archive Of SID



## Evaluating non-parametric supervised classification algorithms in land cover map using Landsat-8 Images

V. Mirzaei Zadeh <sup>1\*</sup>, M. Niknejad <sup>2</sup>, J. Oladi Qadikolaei <sup>3</sup>

1. MSc. Graduated of Forestry, Ilam University

2. PhD. Student of Forestry, Sari University of Agricultural Sciences and Natural Resources

3. Assoc. Prof. College of Natural Resources, Sari University of Agricultural Sciences and Natural Resources

### ARTICLE INFO

#### Article history:

Received 27 December 2014

Accepted 12 July 2015

Available online 16 December 2015

#### Keywords:

Image classification

Land cover maps

Remote sensing

Support Vector Machine

Arakvaz watershed

### ABSTRACT

The aim of this study was to evaluate the efficiency of three support vector machine algorithms, fuzzy decision trees and neural networks for mapping land vegetation map of Arakvaz watershed using OLI sensor of Landsat images (2014). Geometric correction and image pre-processing were utilized to determine the training samples of land vegetation classes for the classification operations. Sample resolution in the vegetation classes has been evaluated using a statistical divergence index. On the next stage, to evaluate the accuracy of algorithms' classification results, ground truth map with the dimensions of 550 m was designed using systematic approach and land vegetation types in the sampling plots were determined. Finally, the efficiency of each classification method was investigated by such criteria as overall accuracy, kappa coefficient, producer accuracy and user accuracy. Comparing the accuracy and kappa coefficient obtained for three categories with a proper band set in comparison with the ground truth map indicates that the Support Vector Machine (SVM) classifier with overall accuracy of 91.26% and kappa coefficient of 0.8731 has had more appropriate results than other algorithms. The results showed that the separation and classification of forest lands with high accuracy have been performed as compared to the other land use classes.

\* Corresponding author e-mail address: [Vahidmirzaei6764@gmail.com](mailto:Vahidmirzaei6764@gmail.com)