



سجش از دور

GIS ایران



سنجش از دور و GIS ایران سال یازدهم، شماره چهارم، زمستان ۱۳۹۸  
Iranian Remote Sensing & GIS Vol.11, No. 4, Winter 2020

۸۳-۱۰۰

## ارائه روشی تلفیقی مبتنی بر الگوریتم‌های طبقه‌بندی پارامتریک و غیر پارامتریک به منظور جداسازی پوشش‌های مختلف در جنگل‌های هیرکانی

محمد سعادت<sup>۱</sup> و رضا شاه‌حسینی<sup>۲\*</sup>

۱. دانشجوی کارشناسی ارشد سنجش از دور، دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، پردیس دانشکده‌های فنی، دانشگاه تهران

۲. استادیار دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، پردیس دانشکده‌های فنی، دانشگاه تهران

تاریخ پذیرش مقاله: ۱۳۹۸/۱۰/۰۷

تاریخ دریافت مقاله: ۱۳۹۷/۱۱/۲۰

### چکیده

تهیه نقشه‌های پوشش اراضی با دقت بالا، همواره یکی از اهداف مهم محققان در زمینه مدیریت اراضی بوده است. هدف از این پژوهش، ارائه روش نوینی جهت تهیه نقشه‌های کاربری اراضی با استفاده از پردازش تصاویر ماهواره‌ای بوده است. به همین منظور، از تصاویر ماهواره لندست ۸، به عنوان تصویر پایه و نقشه مدل رقومی ارتفاعی (DEM)، داده‌های حاصل از تجزیه به عنوان مولفه‌های اصلی و شاخص‌های طیفی جهت استخراج نقشه پوشش اراضی در منطقه مطالعاتی استفاده شد. پس از پیش‌پردازش‌ها و آماده‌سازی داده‌های مورد نیاز، اقدام به تهیه نمونه‌های آموزشی شد. در این پژوهش، نمونه‌های آموزشی در دو بخش به کار گرفته شدند؛ در بخش اول از آنها به عنوان ورودی، جهت طبقه‌بندی تصویر با الگوریتم‌های نظارت شده، حداکثر احتمال و ماشین بردار پشتیبان استفاده شد. در بخش دوم، به منظور طبقه‌بندی با روش درخت تصمیم‌گیری، از این نمونه‌ها برای تعیین محدوده بازتاب طیفی هر پوشش در طیف امواج الکترومغناطیس (باند‌های تصویر، PCA، شاخص‌های طیفی و DEM) استفاده شد. سپس با استفاده از این داده‌ها و شروط دودویی درخت تصمیم‌گیری، هر پوشش مشخص و نقشه پوشش آن استخراج شد. پس از تهیه نقشه‌های ذکر شده، به منظور تلفیق نتایج طبقه‌بندی و حصول دقت بالاتر، از روش حداکثر رای‌گیری به منظور تهیه نقشه تلفیقی جدید پوشش اراضی منطقه استفاده شد. همچنین به منظور ارزیابی دقت نقشه‌های تولیدی، از پارامترهای آماری منتج از ماتریس ابهام شامل دقت کلی، ضریب کاپا، دقت کاربر و دقت تولیدکننده استفاده شد. بر اساس نتایج حاصله، روش تلفیقی با دقت کلی ۹۲/۳۷ درصد و ضریب کاپا ۰/۹۱ دارای بیشترین دقت بوده است. دقت کلی نقشه پوشش روش درخت تصمیم‌گیری، ماشین بردار پشتیبان و حداکثر احتمال نیز به ترتیب ۸۹/۶۱، ۸۸/۰۱ و ۸۷/۶ درصد بوده‌اند. با توجه به اینکه در طبیعت پوشش خالص، به ندرت مشاهده می‌شود و بیشتر پوشش‌ها به صورت ترکیبی وجود دارند، لذا بهتر است از روش‌های نوینی که همه ابعاد پدیده‌ها را پوشش می‌دهند استفاده شود. در این پژوهش، اطلاعات حاصل از طبقه‌بندی نظارت شده و همچنین اطلاعات حاصل از روش منطقی درخت تصمیم‌گیری با یکدیگر تلفیق شده و نتایج حاصله به خوبی، بیانگر بهبود دقت نهایی طبقه‌بندی بودند.

**کلید واژه‌ها:** طبقه‌بندی حداکثر رای‌گیری، پوشش اراضی، الگوریتم‌های طبقه‌بندی، سنجش از دور، لندست ۸

\*نویسنده عهده‌دار مکاتبات: ایران، تهران، خیابان امیرآباد شمالی، پردیس دانشکده‌های فنی دانشگاه تهران، دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، گروه فتوگرامتری و سنجش از دور، تلفن همراه: ۰۹۱۲۲۱۶۱۹۴۹

## ۱- مقدمه

(Y. Shi et al., ۲۰۱۰). سنجش از دور می‌تواند با دقت در تعیین نواحی به ما کمک کند و در تهیه نقشه‌های پوشش اراضی مورد استفاده قرار گیرد. داده‌های ماهواره‌ای، دید مستقیم وسیعی از سطح زمین ارائه می‌کند و باعث می‌شود تا بتوان نقشه پوشش منطقه را تهیه و آن را مورد ارزیابی قرار داد (S. B. Al Rawashdeh, ۲۰۱۰). یکی از روش‌های پرکاربرد استخراج اطلاعات از تصاویر ماهواره‌ای، طبقه‌بندی است که به کاربران امکان تولید اطلاعات مختلف را می‌دهد. برای تفکیک پدیده‌های موضوعی و استخراج دقیق اطلاعات، با توجه به نوع روش طبقه‌بندی، منطقه مورد مطالعه، ویژگی‌های نقاط تعلیمی و نقاط کنترل زمینی نتایج متفاوتی حاصل می‌شود (Sabzghabaei et al., 2015). در این بین، الگوریتم‌های مختلفی بر اساس روش‌های پیکسل پایه، برای تهیه نقشه‌های پوشش اراضی با استفاده از تصاویر ماهواره‌ای تعریف شده‌اند. در روش حداکثر احتمال (MLC<sup>۱</sup>)، طبقه‌بندی بر اساس واریانس و کوواریانس انجام می‌شود و در طبقه‌بندی، پیکسل مورد نظر به طبقه‌ای نسبت داده می‌شود که بیشترین احتمال تعلق پیکسل به آن طبقه وجود دارد (A. J. Sanhouse-García et al., ۲۰۱۶). در روش ماشین بردار پشتیبان (SVM<sup>۲</sup>)، یک مرز تصمیم‌گیری خطی بهینه، برای جدا کردن طبقه‌ها ایجاد می‌شود و با استفاده از همه باندها و یک الگوریتم بهینه‌سازی، نمونه‌هایی که مرزهای طبقه‌ها را تشکیل می‌دهند طبقه بندی می‌شود (Y. Shao et al., ۲۰۱۲). مزیت عمده این روش این است که به تعداد نمونه‌های تعلیمی حساس نیست و برای طبقه‌بندی مناطقی که داده‌های کمی از آنها در دسترس است، بسیار مناسب است (P. Mantero et al., ۲۰۰۵). در روش درخت تصمیم‌گیری (DTC<sup>۳</sup>) تفکیک کلاس پوشش‌ها بر اساس تصمیمات دودویی است و تفکیک پیکسل‌ها به روش سلسله مراتبی و بر اساس شروط منطقی صورت

پوشش اراضی در مفهوم کلی، به آن نوع استفاده از زمین که حاصل دخالت انسان در وضعیت موجود باشد، گفته می‌شود که در برگیرنده تمامی پوشش در بخش‌های مختلف کشاورزی، شهری و روستایی و صنعت می‌شود. به عبارت دیگر، شامل تمام فعالیت‌های موجود در منطقه یا ناحیه مانند یک حوضه آبریز در روی زمین مانند تخصیص اراضی به فعالیت‌های زراعی (دیم و آبی)، مناطق مسکونی، جنگل، مرتع، معدن، تاسیسات صنعتی و همانند آن است (Salajegheh et al., 2011)

دانشمندان علوم طبیعی، اصطلاح پوشش اراضی را به عنوان نشانه فعالیت‌های انسانی از قبیل کشاورزی، جنگلداری و ساخت و ساز که فرآیندهای سطح زمین از جمله بیوژئوشیمی، هیدرولوژی و تنوع زیستی را تغییر می‌دهند، تعریف می‌کنند. پوشش اطلاع از نسبت پوشش‌ها و نحوه تغییرات آن در گذر زمان، یکی از مهمترین موارد در برنامه‌ریزی‌ها است. با اطلاع از نسبت تغییرات پوشش‌ها در گذر زمان می‌توان تغییرات آنی را پیش‌بینی کرده و اقدامات مقتضی را انجام داد (Amirnejad, 2013)

به منظور بررسی پوشش اراضی، می‌بایست در مرحله اول نقشه‌های پوشش اراضی تهیه شود. نقشه‌های پوشش اراضی محققان و سیاست‌گذاران را برای اتخاذ تصمیم صحیح و اصولی جهت مدیریت منابع محیطی یاری می‌کند. مشکل هزینه، جمع‌آوری نمونه‌ها و داده‌ها از یک طرف و زحمت تهیه نقشه‌های پوشش اراضی باعث شده است تا روش‌های نوین دقیق و کم‌هزینه‌تر مورد نیاز باشد. در این بین، سنجش از دور و سامانه اطلاعات مکانی برای تجزیه و تحلیل تغییرات محیط و پوشش سطح زمین ابزار اساسی محسوب می‌شوند. در سال‌های اخیر مطالعه وضعیت پوشش اراضی پیشرفت کرده و روش‌های متنوع جدیدی، ایجاد و برای تحلیل پوشش تغییر پوشش اراضی به کار گرفته شده‌اند که شامل استفاده از سنجش از دور، سامانه اطلاعات مکانی و روش‌های آماری می‌شود (Z. Li et al., ۲۰۰۴) و

1. Maximum Likelihood Classification
2. Support Vector Machine
3. Decision Tree Classifier

است، عمل می‌کند. برچسب‌ها، توسط سخت کردن تصمیم‌های نرم خروجی و با استفاده از انتخاب کننده مقدار حداکثر به دست می‌آید. مزیت مهم این روش این است که می‌توان به طور مستقیم آن را اعمال کرد، حتی در زمانی که شخص متخصص، خود احتمالات اختصاص هر کلاس را به صورت خروجی مشخص نمی‌کند. همچنین مشخص شده است که این روش تلفیقی برای کار بر روی داده‌های سنگین مناسب‌تر است (J. Kittler et al., ۱۹۹۸). از طرف دیگر، با توجه به اینکه، این روش بر اساس حداکثر میزان رای داده شده به هر پیکسل آن را طبقه‌بندی می‌کند، احتمال طبقه‌بندی صحیح‌تر تصویر خروجی در مقایسه با هر یک از الگوریتم‌ها بالا می‌رود؛ لذا انتظار می‌رود دقت نتایج بالاتر رود. از دیگر مزایای این روش نیز می‌توان به قابلیت گسترش آن به هر تعداد تصویر طبقه‌بندی موجود اشاره کرد. همچنین در این روش می‌توان وزن بیشتری به تصویر طبقه‌بندی شده خاصی اختصاص داد (G. James, ۲۰۱۴).

در این پژوهش هدف بر این است که روشی تلفیقی با استفاده از نتایج سایر طبقه‌بندی‌ها ارائه شود تا ضعف هر کدام از آنها پوشانده شده و از مزیت‌های آنها در نتیجه نهایی طبقه‌بندی استفاده شود. با توجه به اینکه در این پژوهش از ویژگی‌های متنوعی نظیر شاخص‌های طیفی، مولفه‌های اصلی، مدل رقومی ارتفاعی و باندهای طیفی سنجنده تصویرگر عملیاتی زمین (OLI) برای اجرای طبقه‌بندی استفاده شده است و هدف پژوهش، مقایسه کارایی روش‌های پارامتریک و غیر پارامتریک طبقه‌بندی پوشش‌های مختلف اراضی جنگلی است، لذا محدودیتی از نظر داده‌های ورودی وجود ندارد و نتایج حاصله به دلیل عدم وجود و یا نقص داده‌های ورودی تحت تاثیر قرار نخواهد گرفت و نتایج، صرفاً بر اساس کارایی الگوریتم‌های طبقه‌بندی حاصل خواهد شد.

می‌گیرد که مناسب ارزیابی مدل‌های پیش بینی چند متغیره است.

تاکنون مطالعات مختلفی در بررسی روش‌های مختلف طبقه بندی صورت گرفته است، اما نتایج متفاوتی حاصل شده است؛ در حالیکه برخی نتایج، بیانگر این بود که الگوریتم حداکثر احتمال، بالاترین دقت را داشته است (C. Li et al., ۲۰۱۴) و (Alizadeh et al., 2016) و (Zandieh, 2015) اما نتایج دیگری نیز گزارش شده است که در آنها الگوریتم ماشین بردار پشتیبان، کارایی بهتری داشته است (Yousofi et al., 2014). همچنین نتایج دیگری نیز گزارش شده که دقت نقشه‌های تولیدی با استفاده از روش شی پایه، نسبت به روش‌های پیکسل پایه بالاتر است و در بین روش‌های پیکسل پایه، دقت روش درخت تصمیم‌گیری بالاتر از سایر روش‌ها بوده است (Fathizad et al., 2015). از طرفی نتایج دیگری نیز گزارش کرده‌اند که استفاده از روش‌های تلفیقی به افزایش ۴ الی ۷ درصدی دقت طبقه‌بندی تصاویر ماهواره‌ای منجر می‌شود (A. Jamil et al., ۲۰۱۸). وجود نتایج مختلف بیانگر این مطلب است که عوامل مختلفی از قبیل منطقه مطالعاتی، نقاط تعلیمی، نوع سنجنده و نوع الگوریتم بر نتایج تاثیرگذار است. روش‌های استفاده شده توسط سایر محققان، همگی بر پایه یک نوع الگوریتم خاص بوده و نتایج متفاوت آنها بیانگر عدم اطمینان کامل از روش‌های قدیمی رایج است. از آنجا که تمامی روش‌های ذکر شده بر پایه بازتاب طیفی پدیده‌ها و مقادیر درجات خاکستری آنها بوده و اختصاص هر پیکسل مجهول از تصویر بر اساس روش آماری مورد استفاده است، لذا مشاهده شده است که دقت نتایج طبقه در هر منطقه، متفاوت است و این مشکل عمده، دقت و کارایی نتایج نهایی را تا حد زیادی دچار مشکل کرده است. روش تلفیقی حداکثر رای گیری (MVC<sup>1</sup>) بر روی برچسب‌های کلاس‌هایی که توسط متخصصان به الگوها اختصاص داده شده

1. Majority Vote Classifiers
2. Operational Land Imager

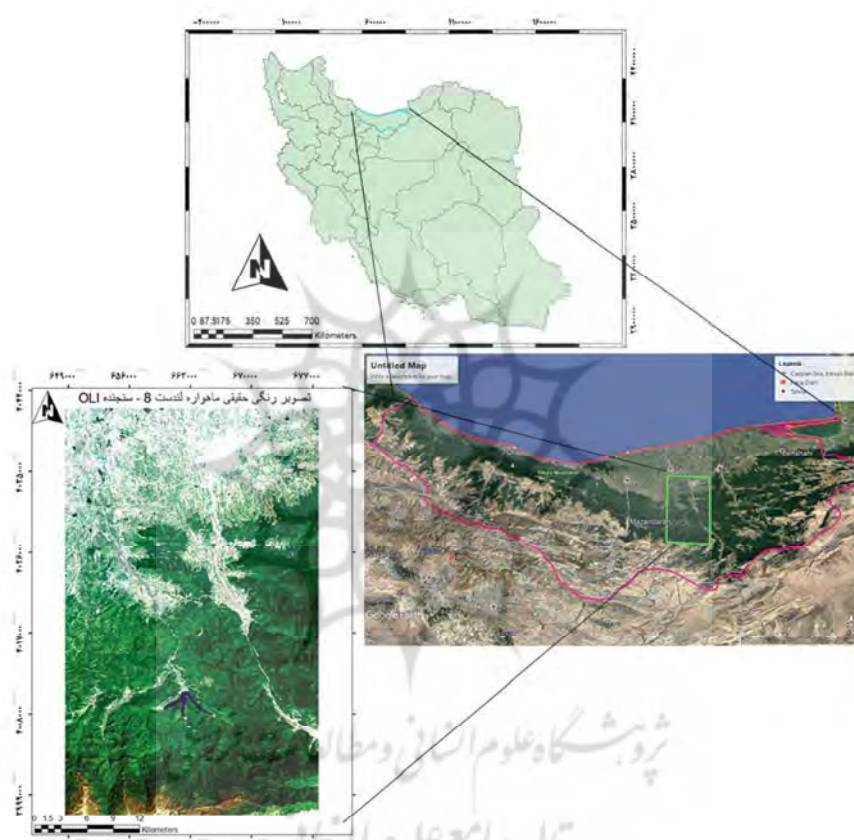
## ۲- مواد و روش‌ها

### ۲-۱- موقعیت منطقه مطالعاتی

در شکل (۱)، منطقه مطالعاتی شامل اراضی جنگل‌های هیرکانی جنوب شهر قائم‌شهر و مناطق اطراف آن نمایش داده شده است. پوشش‌های اصلی این منطقه

شامل جنگل، مرتع، اراضی کشاورزی و باغ، مناطق پوشیده از آب و مناطق ساخته شده است.

در این پژوهش، از تصویر ماهواره لندست ۸ و همچنین مدل رقومی ارتفاعی (DEM<sup>۱</sup>) سنجنده استر<sup>۲</sup> با ابعاد پیکسل ۳۰ متر استفاده شد؛ مشخصات داده‌های مورد استفاده در جدول (۱) ارائه شده است:



شکل ۱. موقعیت جغرافیایی منطقه مورد مطالعه (تصویر رنگی حقیقی ماهواره لندست ۸)

جدول ۱. تصاویر و داده‌های مورد استفاده

نام ماهواره	نام سنجنده	تاریخ برداشت تصویر	ردیف	گذر
Landsat-8	OLI-TIRS	May 2 <sup>nd</sup> , 2018	۱۶۳	۳۵
Terra	ASTER Global DEM	October 11 <sup>th</sup> , 2017	N36	E052

1. Digital Elevation Model
2. ASTER

## ۲-۲- پیش پردازش و تصحیح داده‌ها

تصحیحات رادیومتریکی، شامل آن دسته از تصحیح‌هایی می‌شود که فقط روی درجات خاکستری اعمال شده و با تغییر مقادیر آن‌ها به صورت مجزا (به صورت پیکسل به پیکسل)، سعی در جبران بعضی خطاهای موجود دارند. سنجنده‌های ماهواره‌های لندست، انرژی انعکاسی<sup>۱</sup> خورشیدی را جذب و سپس بسته به نوع سنجنده با تغییر مقیاس این داده‌ها آنها را به مقادیر درجات خاکستری<sup>۲</sup> ۸ یا ۱۶ بیتی تبدیل می‌کنند. به منظور حذف خطاهای رادیومتریکی و اتمسفری تصاویر لندست ۸، طبق روش زیر به مقادیر انعکاسی بالای اتمسفر تبدیل شدند:

### - تبدیل مقادیر درجات خاکستری به انعکاس بالای اتمسفر

این تبدیل از رابطه (۱) صورت می‌گیرد:

$$\rho_{\lambda}' = M_{\rho} Q_{cal} + A_{\rho} \quad \text{رابطه (۱)}$$

که در این رابطه؛  $\rho_{\lambda}'$  بازتاب بالای اتمسفر (بدون انجام تصحیح زاویه تابش خورشیدی)؛  $M_{\rho}$  عامل تغییر مقیاس ضرب شونده که از فایل متا دیتا به دست می‌آید؛  $Q_{cal}$  باند تصویر مورد نظر؛  $A_{\rho}$  عامل تغییر مقیاس اضافه شونده که از فایل متا دیتا به دست می‌آید.

### - تصحیح زاویه تابش خورشید

به منظور تصحیح زاویه تابش خورشید در محاسبه میزان انعکاس بالای اتمسفر از رابطه (۲) استفاده می‌شود:

$$\rho_{\lambda} = \frac{\rho_{\lambda}'}{\cos(\theta_{SZ})} = \frac{\rho_{\lambda}'}{\sin(\theta_{SE})} \quad \text{رابطه (۲)}$$

که در این رابطه؛  $\rho_{\lambda}$  بازتاب بالای اتمسفر (همراه با تصحیح زاویه تابش خورشیدی)؛  $\theta_{SE}$  زاویه محلی ارتفاع خورشید (زاویه ارتفاع خورشید برای مرکز تصویر در فایل متا دیتا بر حسب درجه موجود است)؛  $\theta_{SZ}$  زاویه محلی زینت (اوج) خورشید ( $\theta_{SZ} = 90^{\circ} - \theta_{SE}$ ).

تصاویر ماهواره لندست، عموماً دارای تصحیح هندسی،

کیفیت خوب، ثبت شده منطبق بر پارامترهای مداری و دارای مختصات جغرافیایی هستند و نیاز به تصحیح هندسی ندارند (A. Abdul Qadir et al., ۲۰۱۰).

## ۲-۳- استخراج و انتخاب ویژگی‌های بهینه

معمولاً رفتار طیفی کلاس‌ها در بعضی از بخش‌های طیف الکترومغناطیس با هم شباهت دارد و بنابراین استفاده از این باندها برای طبقه‌بندی، علاوه بر اینکه کمکی به بهبود فرآیند طبقه‌بندی نمی‌کند، بلکه باعث کاهش سرعت و همچنین افزایش تعداد نمونه‌های تعلیمی می‌شود. بنابراین یکی از موارد بسیار مهم در پیش‌پردازش طبقه‌بندی، آنالیز وابستگی باندها و همچنین حذف باندهای اضافه است. در مبحث حذف باندهای اضافه در تصویر لندست ۸، باند شماره ۸ پانکروماتیک، باند شماره ۹ باند بررسی ابرهای سیروس (Cirrus) و باندهای حرارتی ۱۰ و ۱۱ از فرآیند طبقه‌بندی حذف شدند.

شاخص گیاهی تفاضلی نرمال شده ( $NDVI^3$ )، از معروف‌ترین و ساده‌ترین شاخص‌های گیاهی مورد استفاده است و بر اساس دو باند قرمز و مادون قرمز نزدیک، تعریف می‌شود. این شاخص، دارای مقادیر نرمال در دامنه بین -۱ و +۱ است (Fatemi, and Rezaie, 2012). شاخص گیاهی تعدیل‌کننده اثر خاک ( $SAVI^4$ ) نیز با وارد کردن یک پارامتر به فرمول  $NDVI$  سعی در کاهش اثر پس‌زمینه‌های حاکی در مناطق گیاهی دارد.  $L$ ، فاکتور تعدیل‌کننده اثر خاک است که معمولاً با سعی و خطا محاسبه می‌شود (H. Ren et al., ۲۰۱۸).

$$SAVI = (L + 1) \frac{NIR - RED}{NIR + RED - L} \quad \text{رابطه (۳)}$$

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad \text{رابطه (۴)}$$

1. Reflectance
2. Digital Number
3. Normalized Difference Vegetation Index
4. Soil Adjusted Vegetation Index



بوده که جهت طبقه‌بندی تصاویر ماهواره‌ای نیاز به نمونه‌های تعلیمی دارد (G. Mountrakis et al., ۲۰۱۱). تئوری این الگوریتم، توسط وپنیک<sup>۳</sup> به طور کامل بسط یافته و سیستم طبقه‌بندی آن منتج شده از تئوری‌های آماری است (W. N. Vapnik, ۱۹۹۹). با توجه به شکل (۲)، برای محاسبه مرز تصمیم‌گیری دو کلاس تفکیک شده، از روش حاشیه بهینه استفاده می‌شود. محل قرارگیری مرز تصمیم‌گیری با حاشیه بیشینه به گونه‌ای تعیین می‌شود که (۱) تمام نمونه‌های کلاس ۱+ در یک طرف مرز و تمام نمونه‌های کلاس ۱- در طرف دیگر مرز واقع شوند و (۲) فاصله نزدیکترین نمونه‌های آموزشی هر دو کلاس از هم در راستای عمود بر مرز تصمیم‌گیری، تا بیشینه شود. به نمونه‌های آموزشی که شرایط فوق را دارا باشند، بردارهای پشتیبان گفته می‌شود و الگوریتم ماشین بردار پشتیبان تنها از این بردارهای پشتیبان جهت آموزش طبقه‌بندی کننده استفاده می‌کند. کرنل‌های مختلفی در این الگوریتم قابل استفاده است: خطی، چند جمله‌ای، تابع پایه شعاعی و سیگموئیدی (S مانند). در این بین، کرنل تابع پایه شعاعی، در اکثر موارد به عنوان پیش فرض به کار گرفته می‌شود. نمایش ریاضی وار هر یک از کرنل‌ها، به شکل روابط ۶ تا ۹ است.

$$k(x_i, x_j) = x_i^T x_j \quad \text{رابطه (۶) خطی}$$

$$k(x_i, x_j) = (y x_i^T x_j + r)^d, y > 0 \quad \text{رابطه (۷) چندجمله‌ای}$$

$$k(x_i, x_j) = \exp(-y \|x_j - x_i\|^2)^d, y > 0 \quad \text{رابطه (۸) تابع پایه شعاعی}$$

$$k(x_i, x_j) = \tanh(y x_i^T x_j + r) \quad \text{رابطه (۹) سیگموئیدی}$$

یکی از معروف‌ترین روش‌های استخراج عوارض در سنجش از دور، تبدیل مولفه اصلی (PCA<sup>۱</sup>) است. این تبدیل، داده‌های اصلی را به داده‌های کوچکتر که کمتر از داده‌های قبلی به یکدیگر وابسته هستند، تبدیل می‌نماید. بنابراین کاهش تعداد متغیرهای جدید در داده‌های اصلی نهفته است. در حالت کلی، تبدیل مولفه اصلی به دنبال کاهش بعد داده، به منظور پیدا کردن تعداد محدودتری از ترکیب‌های خطی متعامد از داده اصلی با واریانس زیاد است. همیشه به تعداد داده‌های موجود، مولفه اصلی وجود دارد. معمولاً اولین مولفه‌ها، دارای واریانس بیشتری هستند و با حذف بقیه داده‌های با واریانس کمتر، از اطلاعات داده کم نخواهد شد. از آنجایی که واریانس وابسته به مقیاس متغیرهاست، بطور معمول، ابتدا متغیرها را با میانگین صفر و انحراف معیار یک استاندارد می‌نماییم (Hasanloo, and Samadzadegan, 2013).

#### ۲-۴- الگوریتم‌های طبقه‌بندی

##### ۲-۴-۱- الگوریتم حداکثر احتمال

الگوریتم حداکثر احتمال، یکی از معروف‌ترین روش‌های آماری طبقه‌بندی است که جزو روش‌های پیکسل پایه، قرار می‌گیرد. در طبقه‌بندی حداکثر احتمال، کلاسی به پیکسل مورد نظر اختصاص داده می‌شود که بیشترین احتمال تعلق پیکسل به آن کلاس وجود دارد (et ۲۰۱۰). (J. R. Otukei al., در رابطه (۵)، D نشان‌دهنده احتمال یا فاصله وزن دار شده، بردار اندازه‌گیری است،  $M_c$  بر اساس معادله بیزین<sup>۲</sup> و X متعلق به کلاس‌های مشخص شده است.

رابطه (۵)

$$D = \ln(a_c) - [0.5 \ln(|cov_c|)] - [0.5(X - M_c)^T (cov_c - 1)(X - M_c)]$$

مهمترین مزیت این روش، این است که واریانس و کوواریانس درون توزیع کلاس را به حساب آورده و برای داده‌های با توزیع نرمال بهتر عمل می‌کند.

##### ۲-۴-۲- الگوریتم ماشین بردار پشتیبان

الگوریتم ماشین بردار پشتیبان، الگوریتمی غیرپارامتریک

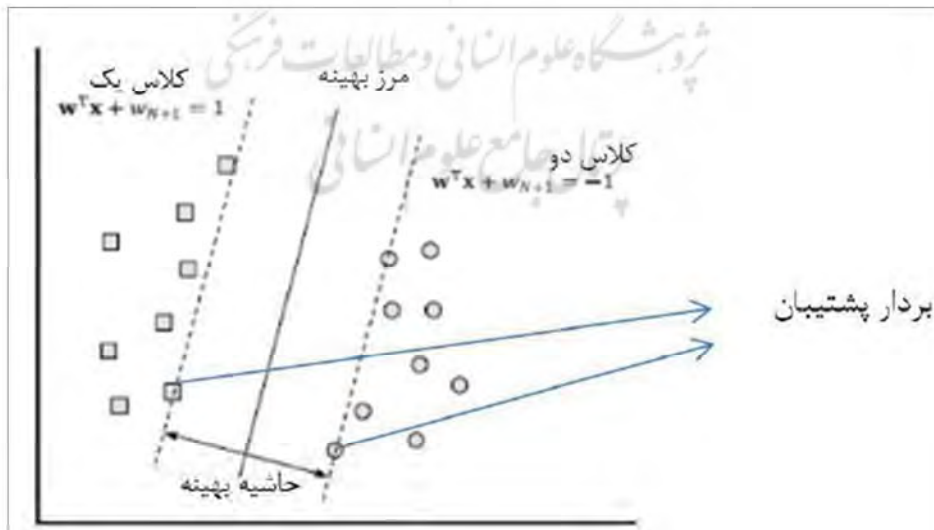
1. Principle Component Analysis
2. Bayesian
3. Vapnik

داده‌ها است. در این روش، هر گره داخلی یا غیر برگ با یک ویژگی مشخص می‌شود. در هر گره داخلی به تعداد پاسخ‌های ممکن شاخه وجود دارد و با پرسش سؤال از گره و جواب بله یا خیر آن (دودویی) آن جدا می‌شوند و هر یک از شاخه‌ها با مقدار جواب آن مشخص می‌شوند. برگ‌های این درخت با یک کلاس که به آن برچسب می‌گوییم و یا یک دسته از جواب‌ها مشخص می‌شوند. بالاترین گره در درخت تصمیم‌گیری را گره ریشه می‌نامند. مدل‌های سلسله مراتبی، بر اساس مجموعه‌ای از تصمیم‌ها و قواعدی است که بر اساس احتمال داده‌های ورودی در هر گره اعمال می‌شوند تا بر مبنای نمونه‌های تعلیمی، به درستی پوشش زمین را طبقه‌بندی کنند. با توجه به شکل (۳)، به منظور تعیین گره‌ها و برگ‌های درخت تصمیم‌گیری، از دو دسته داده ورودی استفاده شد؛ دسته اول داده‌های اولیه (شامل باندهای تصویر و DEM) و دسته دوم داده‌های ثانویه (شامل شاخص‌ها و تبدیل‌ها) هستند. (G. B. Zhu et al., ۲۰۰۶) و (L. et al., ۲۰۱۴) و (Chasmer

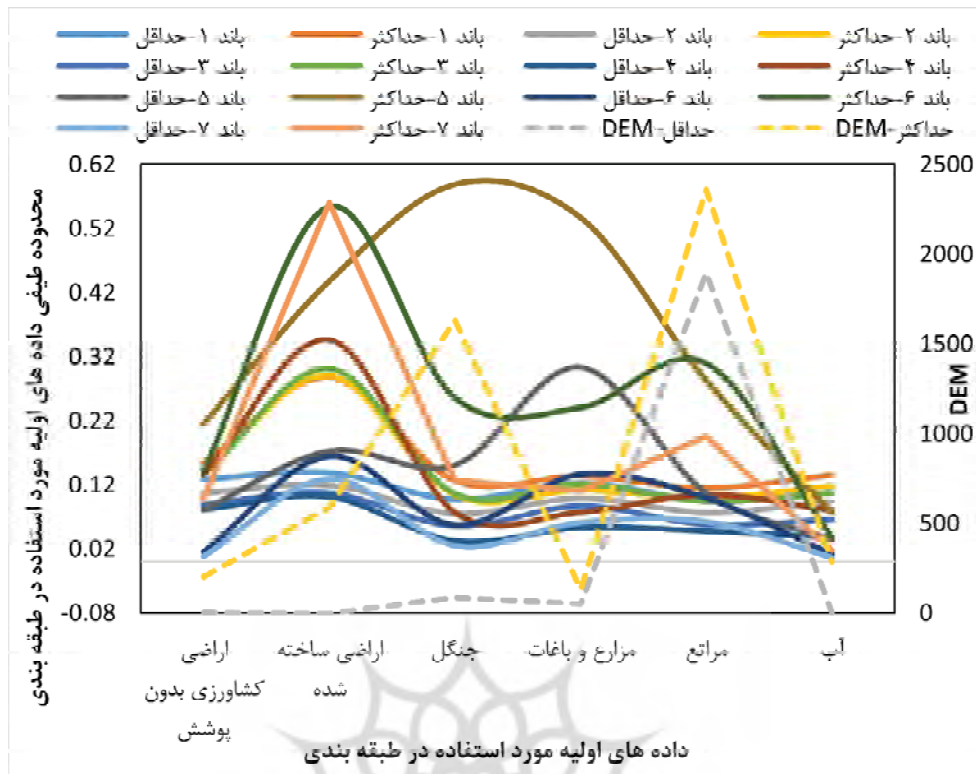
در پایه‌ای ترین حالت، SVM یک طبقه‌بندی کننده خطی باینری است که نمونه‌های آزمایشی داده شده را به یکی از دو حالت ممکن اختصاص می‌دهد. انتخاب تابع کرنل، اغلب نیاز به آنالیز نتایج دارد و مشکل‌ترین مرحله در طبقه‌بندی با این روش است. اگرچه در انتخاب کرنل، گزینه‌های زیادی وجود دارد، اما نتایج محققان نشان داده است که کرنل‌هایی نظیر تابع پایه شعاعی و چندجمله‌ای، در بررسی‌های تجربی نتایج متفاوتی در طبقه‌بندی تصاویر ماهواره‌ای نشان داده‌اند (G. Zhu et al., ۲۰۰۲). در این پژوهش، از کرنل خطی به منظور طبقه‌بندی با روش SVM استفاده شد (S. Hare et al., ۲۰۱۶)، (P. Mather et al., ۲۰۱۶) و (Y. Tang, ۲۰۱۳).

#### ۲-۴-۳- الگوریتم درخت تصمیم‌گیری

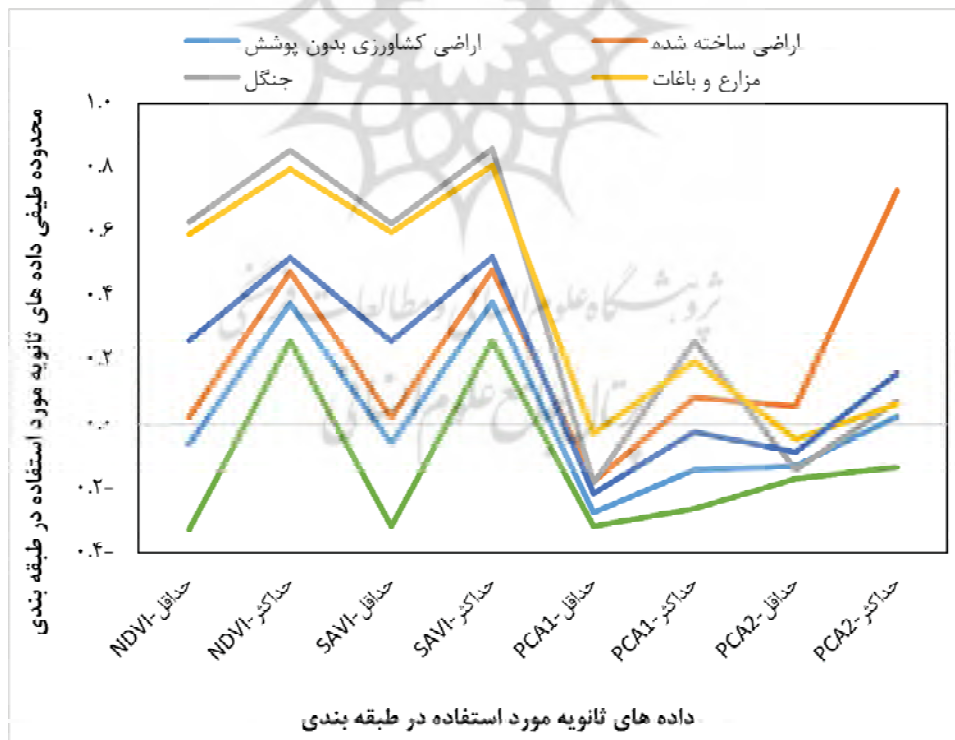
این الگوریتم، بر اساس تفکیک کلاس‌ها بر اساس تصمیمات دودویی، تفکیک پیکسل‌ها به روش سلسله مراتبی، تصمیم‌گیری بر اساس شروط منطقی، رشد درخت از ریشه به برگ و عدم نیاز به دانش اولیه از



شکل ۲. مرز بین بهینه بین کلاس‌ها و بردارهای پشتیبان



(الف)



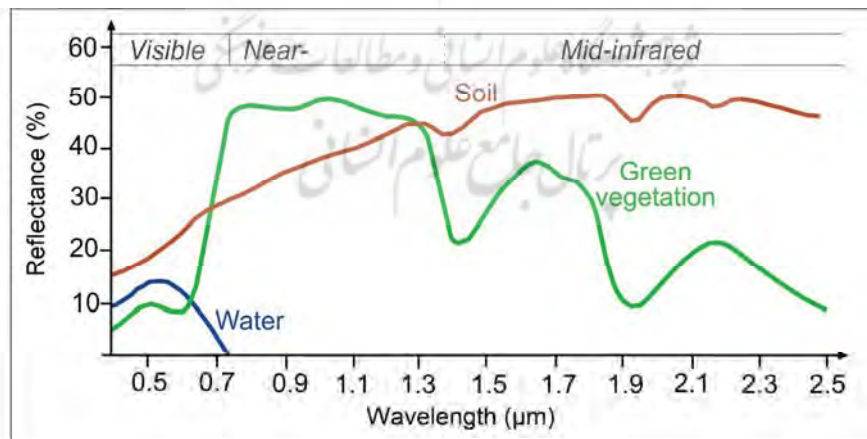
(ب)

شکل ۳. (الف) محدوده بازتاب طیفی استخراج شده از باندهای تصویر ماهواره لندست و (ب) محدوده بازتاب طیفی شاخص های گیاهی و مقادیر بازتاب طیفی حاصل از تبدیل PCA در هر یک از پوشش های اراضی



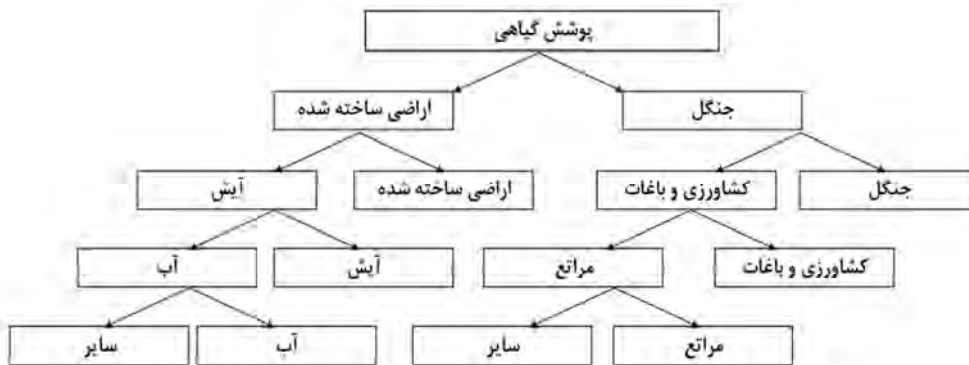
که بیشترین تمایز بین پوشش گیاهی و خاک لخت در طیف طول موج SWIR است. در این طیف طول موج پوشش گیاهی به نسبت خاک لخت، بازتاب به مراتب کمتری داشته است و قابلیت تمایز این دو پدیده را بالا برده است. همچنین طول موج‌های سبز و قرمز نیز اختلاف بازتاب طیفی محسوسی برای تفکیک پوشش گیاهی از خاک لخت مشاهده شده است اما این اختلاف به میزان طیف طول موج SWIR نمی‌باشد. با این حال به دلیل افزایش کارایی مدل پیشنهادی جهت حصول طبقه‌بندی مناسب، از این دو طیف طول موج نیز استفاده شد. همچنین با توجه به نمونه‌های تهیه شده از تصویر و ترسیم منحنی بازتاب طیفی پدیده‌ها، مشخص شد که مولفه دوم حاصل از PCA برای جداسازی سایر پوشش‌ها مناسب‌تر است. با توجه به اینکه شباهت طیفی بالایی بین پوشش‌های جنگل، مزارع و باغات و مراتع وجود داشت، تصویر DEM نقش بسیار مهمی در جداسازی پوشش‌ها ایفا کرد زیرا در منطقه مورد مطالعه، جنگل‌ها در ارتفاعات بالا و مزارع و باغات در دشت قرار داشتند. مراحل اجرای درخت تصمیم‌گیری در شکل ۵، ارائه شده است.

در این روش به منظور جداسازی طیفی کلاس پوشش گیاهی از سایر پوشش‌های اراضی، از شاخص‌های طیفی SAVI و NDVI، مدل رقومی زمین (DEM)، باند مادون قرمز نزدیک ( $NIR^1$ ) و مولفه اول PCA استفاده گردید. با توجه به اینکه در طیف طول موج NIR، پوشش گیاهی دارای بیشترین بازتاب طیفی بوده و با سایر پدیده‌های سطح زمین بیشترین تمایز را دارد و در طیف طول موج قرمز بیشترین جذب توسط پوشش گیاهی صورت می‌پذیرد، لذا چه در شاخص‌های گیاهی و چه در محاسبات طیفی مستقل، از اختلاف طیف طول موج NIR و قرمز به منظور جداسازی پوشش گیاهی از سایر پدیده‌ها استفاده می‌شود (Jr. J. Rouse et al., ۱۹۷۴). از طرف دیگر با توجه به نمونه‌های تهیه شده و ترسیم منحنی بازتاب طیفی پدیده‌ها در تصویر، مشخص شد مولفه اول حاصل از PCA برای جداسازی پوشش‌های گیاهی مختلف مناسب‌تر است. علاوه بر موارد مذکور، از باندهای قرمز، مادون قرمز طول موج کوتاه ( $SWIR^2$ )، سبز و PCA حاصل از مولفه دوم نیز برای جداسازی سایر پوشش‌ها استفاده شد. با توجه به منحنی بازتاب طیفی پدیده‌ها در شکل (۴)، در طول موج‌های مختلف مشخص شد



شکل ۴. منحنی بازتاب طیفی پدیده‌ها [۳۵]

1. Near Infrared
2. Shortwave Infrared



شکل ۵. مراحل اجرای درخت تصمیم‌گیری

در این رابطه مقدار  $\omega_1, \dots, \omega_B$  و وزن‌ها هستند که مجموع آنها برابر با ۱ بوده و  $I(\cdot)$  یک تابع نشانگر است. اگر وزن‌ها به صورت  $1/B$  تنظیم شوند، در این شرایط حالت  $h_1, h_2, \dots, h_B$  به صورت رابطه ۹ خواهد بود.

**۲-۵- ارزیابی دقت الگوریتم‌های طبقه‌بندی کننده**  
ارزیابی خطا و برآورد دقت طبقه‌بندی، معمولاً بر اساس پارامترهای آماری است که از ماتریس خطا استخراج می‌شوند. ماتریس خطا که ماتریس ابهام نیز نامیده می‌شود، حاصل مقایسه پیکسل به پیکسل، پیکسل‌های معلوم (در واقعیت زمینی) با پیکسل‌های متناظر در نتایج طبقه‌بندی است. در این پژوهش، به منظور ارزیابی دقت طبقه‌بندی‌های انجام گرفته از ماتریس ابهام و پارامترهای منتج از آن شامل دقت کلی، ضریب کاپا، دقت کاربر، دقت تولیدکننده، خطای حذف ( $O_e$ ) و اضافه ( $C_e$ ) استفاده شد.

### ۳- روش تحقیق و پیاده‌سازی

پس از انجام مراحل آماده‌سازی و پیش پردازش‌های لازم، تصویر لندست ۸، به همراه شاخص NDVI، مولفه اول و دوم حاصل از PCA؛ بیشترین حجم داده‌ها در مولفه اول و سپس مولفه دوم ذخیره شده است

1. Omission
2. Commission

هدف از شکل (۵)، نحوه قرارگیری Node ها و طراحی درخت تصمیم‌گیری است و در هر Node با اعمال حد آستانه بر روی ویژگی‌های مورد استفاده برای هر کلاس پوشش زمین، به جداسازی این کلاس‌ها پرداخته می‌شود. در مرحله اول پوشش گیاهی از سایر پوشش‌ها جدا شده و سپس با توجه به تراکم پوشش گیاهی و مقادیر بازتاب طیفی آن در هر یک از شاخص‌های گیاهی و باندهای تصویر، حد آستانه‌ای برای پوشش‌های مختلف دارای پوشش گیاهی تعریف شده است. حد آستانه نیز بر اساس نمونه‌برداری از پیکسل‌ها و میانگین‌گیری صورت گرفته است.

### ۲-۴-۴- روش تلفیقی حداکثر رای‌گیری

در این روش، اساس کار تلفیق نتایج چند الگوریتم طبقه‌بندی به منظور حصول تصویر طبقه‌بندی شده‌ای مبتنی بر الگوریتم‌های ورودی است. روش مرسوم برای تلفیق نتایج، چند تصویر طبقه‌بندی شده به صورت زیر است:

$$C(X) = \text{mode}\{h_1(X), h_2(X), h_3(X)\} \quad (10)$$

به بیان دیگر، در این رابطه در هر مقدار  $X$ ، به کلاسی تعلق می‌گیرد که بیشترین رای را دریافت کرده است. به این روش طبقه‌بندی، طبقه‌بندی حداکثر رای‌گیری می‌گویند. رابطه نهایی این روش به صورت زیر بیان می‌شود.

$$C(X) = \arg \max \sum_{j=1}^B \omega_j I(h_j(X) = i) \quad (11)$$

با توجه به هدف تحقیق و همچنین پدیده‌های غالب موجود در منطقه، هفت نوع پوشش اراضی در منطقه مطالعاتی با توجه به جدول (۳) تفکیک و شناسایی شدند. به منظور طبقه‌بندی بر اساس روش درخت تصمیم‌گیری و از آنجا که این روش، یک روش دودویی بر اساس شروط منطقی است، ابتدا نمونه‌های تعلیمی برای هر پوشش انتخاب و سپس بر اساس محدوده بازتاب طیفی هر پدیده در طیف طول موج‌های مختلف، حدود آستانه برای هر پوشش تعیین شد. سپس با استفاده از این محدوده‌های طیفی، شروط منطقی برای هر پوشش استفاده شده و شاخه‌های درخت تصمیم‌گیری شکل گرفتند. به منظور بالابردن دقت طبقه‌بندی از داده‌های جانبی نظیر شاخص‌های طیفی و نقشه رقومی ارتفاع همراه با مولفه‌های PCA نیز استفاده شد و سپس طبقه‌بندی صورت پذیرفت.

(M. M. Mather, ۱۹۹۹)؛ و تصویر DEM سنجنده استر (با پیکسل سایز یکسان با تصویر لندست) وارد مراحل پردازش و طبقه‌بندی شده و سپس به منظور تعیین نمونه‌های تعلیمی و استخراج محدوده بازتاب طیفی پدیده‌ها در طیف طول موج‌های مختلف، با توجه به نقاط واقعیت زمینی و تصاویر گوگل ارث، اقدام به تعیین نمونه‌های تعلیمی و نمونه‌های واقعیت زمینی شد. نمونه‌های تعلیمی، پیکسل‌هایی هستند که معرف هر پوشش بر روی تصویر بوده و سایر پیکسل‌های مجهول با توجه به آنها به هر کلاس اختصاص داده می‌شوند، در حالیکه نمونه‌های واقعیت زمینی پیکسل‌هایی هستند که از نوع پوشش آنها اطمینان کامل وجود داشته و به منظور تعیین دقت طبقه‌بندی مورد استفاده قرار می‌گیرند. تعداد پیکسل‌هایی که برای هر پوشش به عنوان نمونه‌های تعلیمی استفاده شده است، در جدول (۲) ارائه شده‌اند.

جدول ۲. تعداد پیکسل‌های انتخاب شده به عنوان نمونه‌های آموزشی جهت طبقه‌بندی

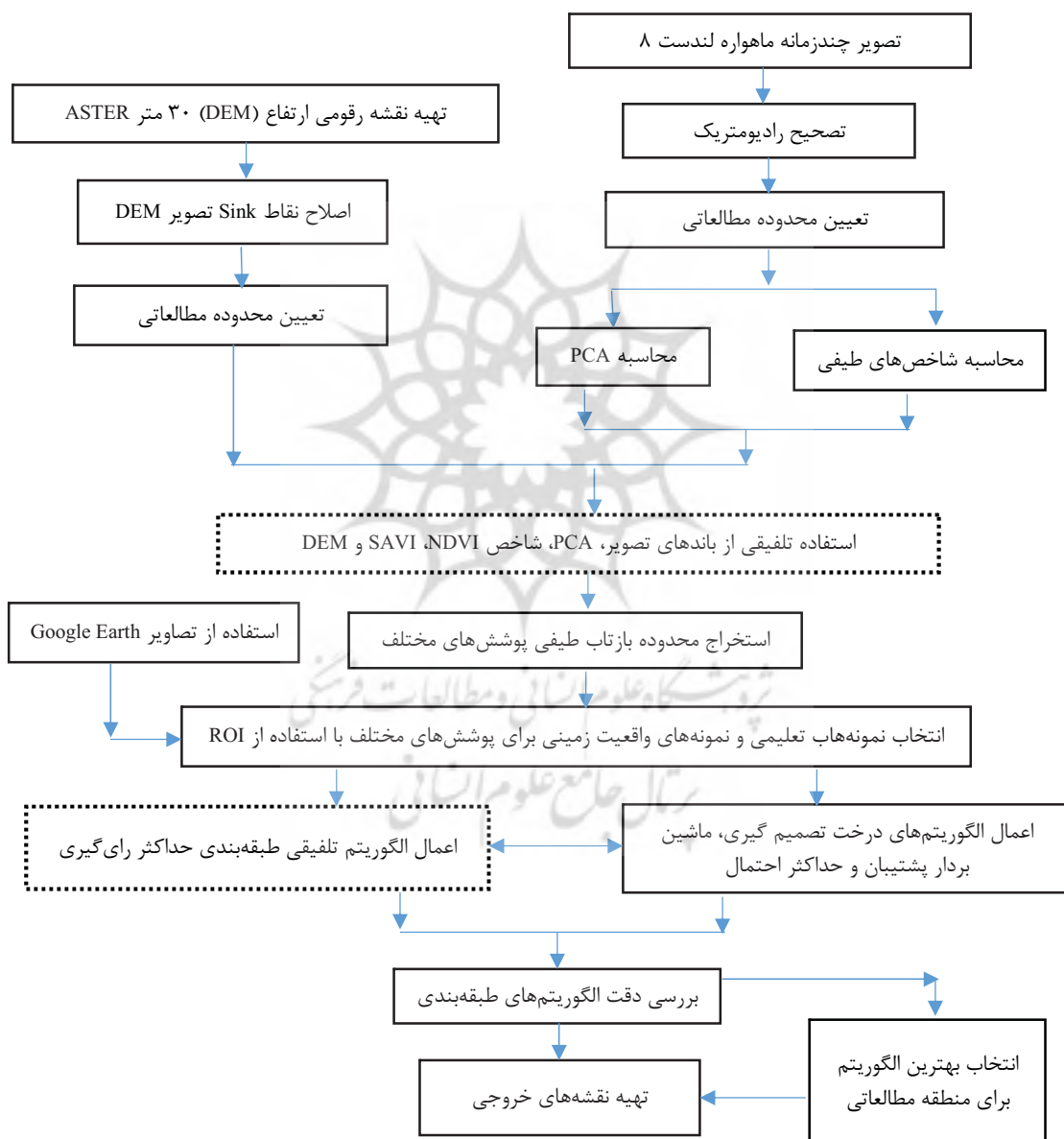
شماره	کلاس پوشش زمین	تعداد پیکسل
۱	جنگل	۷۷۴۸۴
۲	مرتع	۱۸۱۴
۳	اراضی ساخته شده	۲۳۸۰
۴	کشاورزی و باغ	۳۲۶۳
۵	آب	۸۵۲
۶	زمین‌ایش	۵۱۲۷
۷	سایر پوشش‌ها	۹۴۰

جدول ۳. مشخصات هر یک از پوشش‌های تفکیک شده

پوشش‌ها	مشخصات
جنگل	جنگل‌های هیرکانی
مرتع	شامل مراتع موجود بر روی تپه‌ها و دامنه کوه‌ها
اراضی ساخته شده	شامل مناطق مسکونی شهری، روستایی و نواحی صنعتی
کشاورزی و باغ	اراضی کشاورزی دارای پوشش گیاهی و باغات
آب	رودخانه، دریاچه مصنوعی، دریاچه پشت سد و حوضچه‌ها شهری
زمین‌ایش	اراضی کشاورزی شخم زده شده یا ایش
سایر پوشش‌ها	اراضی متفرقه، بیرون زدگی سنگی
طبقه‌بندی نشده	شامل تمام پیکسل‌های طبقه‌بندی نشده

تلفیقی حداکثر رای گیری به منظور کسب بهترین نتیجه استفاده شد. روش حداکثر رای گیری با استفاده از نتایج سایر طبقه بندی ها، سعی در پوشش ضعف هر یک از سایر طبقه بندی ها و استفاده از نتایج بهینه هر یک از آنها در نتیجه نهایی داشته است. دیاگرام مراحل پژوهش در شکل (۶) ارائه شده است.

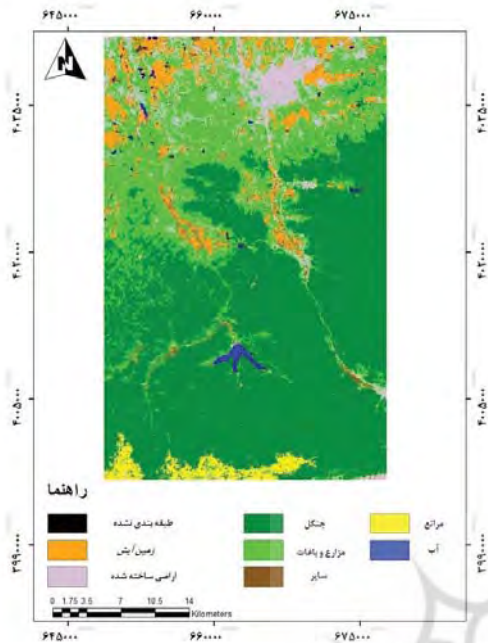
به منظور طبقه بندی با استفاده از الگوریتم های حداکثر احتمال و ماشین بردار پشتیبان نیز از نمونه های تعلیمی تهیه شده برای الگوریتم درخت تصمیم گیری به عنوان نمونه های تعلیمی استفاده شد تا شرایط انجام طبقه بندی برای همه الگوریتم های یکسان باشد. در انتها به منظور استخراج بهترین نتیجه از روش



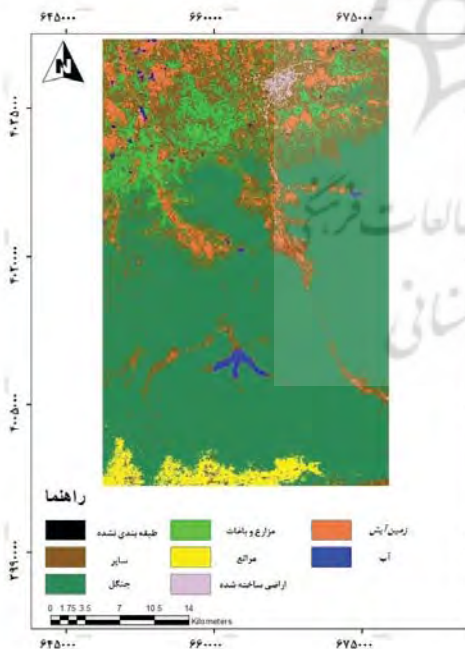
شکل ۶. نمودار جریان‌ی مراحل پردازش

#### ۴- نتایج تجربی

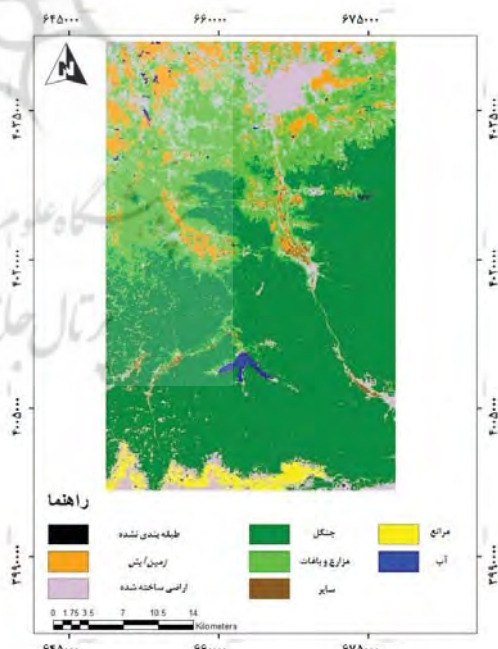
پس از طبقه‌بندی تصاویر ماهواره‌ای و تهیه نقشه‌های پوشش اراضی، بهترین روش طبقه‌بندی در منطقه مطالعاتی با استفاده از دقت‌های به‌دست آمده، تعیین شد. با توجه به نتایج حاصل از طبقه‌بندی در جدول (۴) و همچنین دقت کلی و ضریب کاپا مشخص شد روش حداکثر رای‌گیری با دقت کلی ۹۳/۳۷ درصد و ضریب کاپا ۰/۹۱ به مراتب بهتر از سایر الگوریتم‌ها عمل کرده است. در مرتبه بعدی طبقه‌بندی درخت تصمیم‌گیری، ماشین بردار پشتیبان و در آخر الگوریتم حداکثر احتمال قرار می‌گیرند. با توجه به اینکه ضریب کاپا همه روش‌ها بیش از ۰/۷۵ بوده است لذا می‌توان با اطمینان گفت نقش شانس در طبقه‌بندی پیکسل‌های مجهول بسیار پایین بوده است و طبقه‌بندی بر اساس نمونه‌های تعلیمی انجام شده است. نقشه پوشش اراضی تهیه شده با الگوریتم حداکثر احتمال، ماشین بردار پشتیبان، درخت تصمیم‌گیری و حداکثر رای‌گیری به ترتیب در شکل‌های (۱۰-۷) نمایش داده شده است.



شکل ۸. نقشه پوشش اراضی تهیه شده با الگوریتم ماشین بردار پشتیبان



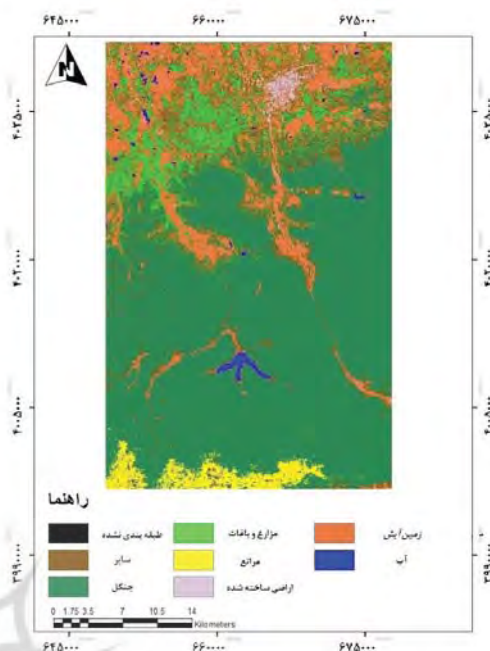
شکل ۹. نقشه پوشش اراضی تهیه شده با الگوریتم درخت تصمیم‌گیری



شکل ۷. نقشه پوشش اراضی تهیه شده با الگوریتم حداکثر احتمال



با توجه به دقت کاربر و خطای اضافه در جداول (۵) و (۶) مشخص شد که بیشترین دقت در هر چهار الگوریتم طبقه‌بندی، به ترتیب به پوشش‌های آب و مرتع بوده است؛ درحالی‌که در این بین کمترین دقت به ترتیب، مربوط به سایر پوشش‌ها و اراضی ساخته‌شده، بوده است. علت این امر این است که این دو پوشش، ترکیبی از چند نوع پوشش بوده و در نتیجه شباهت طیفی بالای آنها در طیف الکترومغناطیس با سایر پوشش‌ها منجر به کاهش دقت طبقه‌بندی شده است. با این وجود در روش تلفیقی، حداکثر رای‌گیری دقت کاربر برای این دو پوشش اراضی بیش از ۶۰ درصد بوده، درحالی‌که در سایر طبقه‌بندی‌ها، دقت کمتر از ۲۰ درصد مشاهده شده است. همچنین در روش حداکثر رای‌گیری، دقت پنج پوشش اراضی دیگر بالای ۹۵ درصد شده است که نسبت به سه روش دیگر به مراتب بهتر بوده است. در حقیقت دقت کاربر احتمال طبقه‌بندی یک کلاس خاص در نقشه، مطابق با همان کلاس در زمین است و در ماتریس خطا نسبت پیکسل‌های صحیح به مجموع پیکسل‌های یک سطر است.



شکل ۱۰. نقشه پوشش اراضی تهیه شده با الگوریتم حداکثر رای‌گیری

#### جدول ۴. دقت کلی و ضریب کاپا طبقه‌بندی

الگوریتم‌های SVM، DTC، MLC و حداکثر رای‌گیری

روش حداکثر رای‌گیری	درخت تصمیم‌گیری	ماشین بردار پشتیبان	حداکثر احتمال	دقت کلی (%)	ضریب کاپا
۹۳/۳۷	۸۹/۶۱	۸۸/۰۱	۸۷/۶		
۰/۹۱	۰/۸۴	۰/۸۲	۰/۸۲		

#### جدول ۵. دقت تولیدکننده و دقت کاربر طبقه‌بندی الگوریتم‌های SVM، DTC، MLC و روش حداکثر رای‌گیری

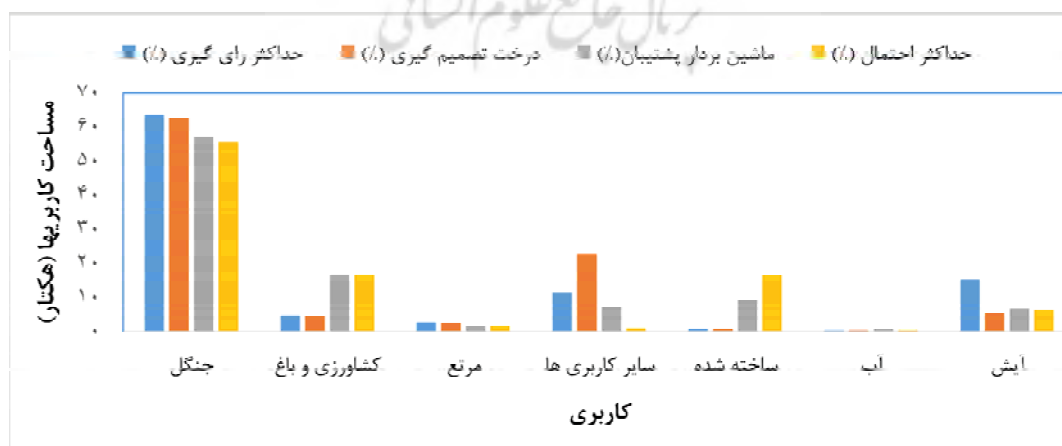
MVC	SVM	DTC	MLC	MVC	SVM	DTC	MLC	
۹۷.۰۶	۹۶.۲۵	۹۴.۱۹	۹۵.۹۹	۹۶.۸۵	۹۷.۰۵	۹۷.۲۴	۹۶.۳۷	جنگل
۹۳.۹۶	۹۷.۵۳	۹۳.۱۶	۹۴.۹۵	۹۱.۸	۷۳.۲۲	۷۸.۱۳	۷۵.۲	کشاورزی و باغ
۹۷.۱۷	۹۶.۷۷	۹۴.۸۸	۹۷.۲۶	۹۵.۸۵	۷۰.۱۵	۴۸.۸	۳۳.۹۵	مرتع
۶۶.۲۴	۹۳.۶۹	۹۶.۳۹	۹۷.۴۸	۹۶.۰۵	۹۴.۰۵	۹۵.۷۴	۹۷.۳	زمین‌ایش
۹۴.۰۸	۹۴.۹۳	۷۲.۶۴	۷۹.۴۷	۷۴.۴۴	۳۸.۲۵	۹۷.۴۳	۹۶.۲۲	اراضی ساخته شده
۹۷.۳۶	۹۶	۹۷.۱۹	۹۷.۹۳	۸۹.۵۷	۹۳.۰۲	۹۷.۸۵	۹۲.۹۵	آب
۶۴.۹۲	۱۸.۴۶	۹.۱۳	۱۲.۳۲	۶۷.۶۴	۸۳.۴۱	۱۵.۳	۴۶.۴۶	سایر

ماشین بردار پشتیبان، دارای بیشترین دقت در بین سایر پوشش‌ها بوده‌اند، اما در پوشش مرتع، روش تلفیقی حداکثر رای‌گیری دارای بیشترین دقت و روش حداکثر احتمال دارای کمترین دقت در بین سایر پوشش‌ها بوده است. در واقع دقت تولیدکننده احتمال اینکه تولیدکننده پیکسلی را به کلاس خاص نسبت داده باشد در صورتی که کلاس واقعی آن مشخص باشد را بیان می‌کند و در ماتریس خطا، نسبت پیکسل‌های صحیح به مجموع پیکسل‌های یک ستون است. در کل با توجه به نتایج مشخص شد، دقت تولید کننده روش تلفیقی بجز در پوشش اراضی ساخته شده، آب و سایر در سایر پوشش‌ها بالاتر بوده است. همچنین با توجه به دقت کاربر مشخص شد که به جز پوشش زمین، شش پوشش اراضی دیگر، دقت بالاتری نسبت به سایر روش‌ها داشته‌اند.

همچنین با توجه به دقت تولیدکننده و خطای حذف در جداول (۵) و (۶) مشخص شد که بیشترین دقت در هر چهار روش طبقه‌بندی به ترتیب مربوط به پوشش جنگل و آب بوده است. با توجه به نتایج، به جز روش حداکثر احتمال، سایر روش‌های طبقه‌بندی، پوشش جنگل را با دقت بالایی طبقه‌بندی کرده‌اند. به نظر می‌رسد علت این امر به دلیل محدودیت روش آماری طبقه‌بندی حداکثر احتمال است که شباهت طیفی پوشش‌های زمینی مانع از عملکرد بهتر آن شده است. همچنین در بین روش‌های ذکر شده، پوشش آب توسط روش درخت تصمیم‌گیری با بالاترین دقت طبقه‌بندی شده است. همچنین در بین روش‌های فوق الذکر، پوشش‌های مرتع و سایر، پایین‌ترین دقت‌ها را داشته‌اند. با توجه به نتایج مشخص شد که در پوشش سایر اراضی، روش درخت تصمیم‌گیری کمترین دقت و

جدول ۶. خطای حذف و خطای اضافه طبقه‌بندی الگوریتم‌های SVM، DTC، MLC و روش حداکثر رای‌گیری

Commission(%)				Omission(%)				
MVC	SVM	DTC	MLC	MVC	SVM	DTC	MLC	
۲.۹۴	۳.۷۵	۵.۸۱	۴.۰۱	۳.۱۵	۲.۹۵	۲.۷۶	۳.۶۳	جنگل
۶.۰۴	۲.۴۷	۶.۸۴	۵.۰۵	۸.۲	۲۶.۷۸	۲۱.۸۷	۲۴.۸	کشاورزی و باغ
۲.۸۳	۳.۲۳	۵.۱۲	۲.۷۴	۴.۱۵	۲۹.۸۵	۵۱.۲	۶۶.۰۵	مرتع
۳۳.۷۶	۶.۳۱	۳.۶۱	۲.۵۲	۳.۹۵	۵.۹۵	۴.۲۶	۲.۷	زمین ایش
۵.۹۲	۵.۰۷	۲۷.۳۶	۲۰.۵۳	۲۵.۵۶	۶۱.۷۵	۲.۵۷	۳.۷۸	اراضی ساخته شده
۲.۶۴	۴	۲.۸۱	۲.۰۷	۱۰.۴۳	۶.۹۸	۲.۱۵	۷.۰۵	آب
۳۵.۰۸	۸۱.۵۴	۹۰.۸۷	۸۷.۶۸	۳۲.۳۶	۱۶.۵۹	۸۴.۷	۵۳.۵۴	سایر



شکل ۱۱. نمودار تغییرات درصد مساحت هر یک از پوشش‌ها با توجه به روش‌های طبقه‌بندی

اراضی شد. هر یک از این الگوریتم‌ها در تحقیقات فراوانی مورد بررسی قرار گرفته و نتایج متفاوتی ارائه داده‌اند. نتایج پژوهش حاضر، نشان دادند که الگوریتم درخت تصمیم‌گیری بهترین نتایج را ارائه کرده است. در پژوهشی در جنگل‌های زاگرس در استان ایلام و با استفاده از سنجنده ETM+ گزارش شد بین الگوریتم‌های شبکه مصنوعی آرت‌مپ فازی، حداکثر احتمال و درخت تصمیم‌گیری، الگوریتم درخت تصمیم‌گیری دارای بیشترین دقت و ضریب کاپا بوده است و از الگوریتم‌های دیگر بهتر عمل کرده است (آرخی، ۱۳۹۱)، اما در پژوهش دیگری و با استفاده از تصویر سنجنده OLI گزارش شد که در مناطق نیمه خشک حوضه آبخیز دشت ملایر، الگوریتم حداکثر احتمال نتایج مطلوب‌تری ارائه داده است. همچنین در تحقیق دیگری که با استفاده از تصویر سنجنده OLI صورت گرفت، گزارش شد که در حوضه آبخیز رود کرج، این الگوریتم حداکثر احتمال بوده است که بهترین نتیجه را داده است. اما در گزارش دیگری نشان داده شد که استفاده از روش‌های تلفیقی به جای روش‌های ذکر شده می‌تواند منجر به افزایش ۴ الی ۷ درصدی دقت طبقه‌بندی تصاویر ماهواره‌ای شده است. با توجه به این نکات در این پژوهش، ابتدا طیف متنوعی از داده‌ها در کنار تصاویر ماهواره‌ای انتخاب و تهیه شده، سپس نقشه‌های پوشش اراضی با استفاده از الگوریتم‌های پرکاربرد طبقه‌بندی تصاویر ماهواره‌ای تهیه شدند. پس از آن با استفاده از نتایج حاصله از الگوریتم‌های ذکر شده، با روش تلفیقی حداکثر رای‌گیری اقدام به تهیه نقشه پوشش اراضی جدید شد. نتایج دقت نقشه‌های تولیدی نشان داد که روش تلفیقی با بهره‌گیری از مزایای سایر روش‌های طبقه‌بندی، نقشه‌ای دقیق‌تر ارائه داده است و می‌توان از آن به عنوان روش نوینی جهت تهیه نقشه‌های پوشش اراضی استفاده کرد.

بر اساس نتایج موجود در شکل (۱۱)، پوشش جنگل، مراتع و زمین‌های آیش در الگوریتم حداکثر رای‌گیری نسبت به سایر الگوریتم‌ها درصد مساحت بیشتری را به خود نسبت داده است؛ درحالی‌که در الگوریتم ماشین بردار پشتیبان، پوشش کشاورزی و باغات؛ بیشترین درصد مساحت را به خود اختصاص داده است. در بین چهار روش طبقه‌بندی، الگوریتم حداکثر احتمال نسبت به سایر الگوریتم‌ها درصد مساحت پوشش جنگل را کمتر نشان داده است و از طرفی با توجه به جدول (۵)، دقت تولیدکننده کمتری نیز داشته است (هر چند دقت کاربر آن‌اینگونه نبوده است). در زمین‌های آیش، الگوریتم حداکثر رای‌گیری بیشترین مساحت را با اختلاف نسبت به سایر روش‌ها تفکیک کرده است. در پوشش آب، تقریباً تمام روش‌ها نتایج یکسانی ارائه داده‌اند و درصد مساحت آب در آنها بسیار نزدیک است. در بخش سایر پوشش‌ها، درخت تصمیم‌گیری در مقایسه با سایر روش‌ها، مساحت بیشتری را به خود اختصاص داده است که این میزان نسبت به سایر روش‌ها نسبتاً مقدار زیادی است.

## ۵- بحث و نتیجه‌گیری

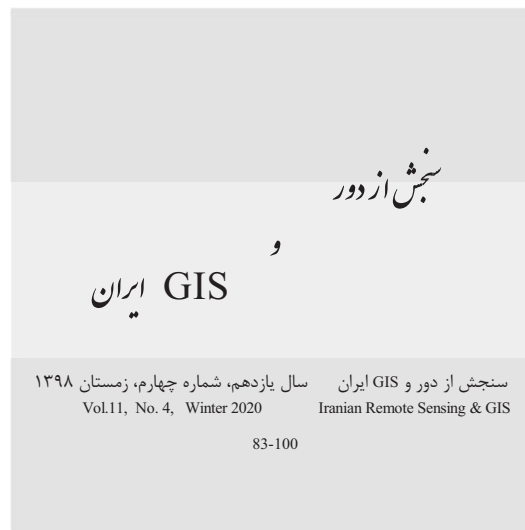
آگاهی از چگونگی تغییرات و روند فعالیت پوشش گیاهی می‌تواند به برنامه‌ریزی‌های راهبردی و کلان مرتبط و اثرگذار بر پوشش گیاهی کمک شایانی کند. با توجه به اهمیت حفظ جنگل‌های شمال در سطح ملی و بین الملل به عنوان تنها گونه جنگل‌های هیرکانی، بررسی و آگاهی از تغییرات این جنگل‌ها بسیار حائز اهمیت است. به منظور آگاهی از این مهم و تهیه نقشه‌های پوشش اراضی، تاکنون روش‌های سنتی و جدید متنوعی استفاده شده است. یکی از بهترین روش‌ها، استفاده از داده‌های دورسنجی و قابلیت بازتاب طیفی پدیده‌ها در طول موج‌های مختلف است. در این پژوهش، اقدام به بررسی چهار روش پیشرفته طبقه‌بندی تصاویر ماهواره‌ای برای تهیه نقشه پوشش

۶- منابع

- Salajegheh, A., Razavizade, S., Khorasani, N., Hamidifar, M., & Salajegheh, S. 2011, **Land Use Changes and its Effects on River Water Quality (Case Study: Karkheh Watershed)**, Journal of Environmental Studies, Vol. 58, pp. 81-86.
- Helming, K., Perez-Soba, M., & Tabbush, P., 2008, **Sustainability Impact Assessment of Land Use Changes**. USA: Springer-Verlag Berlin Heidelberg Press.
- Amirnejad, H., 2013, **Factors Affecting Farmers' Willingness to Changing Land Use in Mazandaran Province**, Journal of Agricultural Economics Research, Vol. 5(4), pp. 87-106.
- Li, Z., Li, X., Wang, Y., Ma, A. & Wang, J., 2004, **Land-use change analysis in Yulin prefecture, northwestern China using remote sensing and GIS**, International Journal of Remote Sensing, Vol. 51, pp. 23-55.
- Shi, Y., Wang, R., Fan, L., Li, J., & Yang, D., 2010, **Analysis on Land-use Change and Its Demographic Factors in the Original-stream Watershed of Tarim River Based on GIS and Statistic**, **Procedia Environmental Sciences**, Vol. 2, pp. 175-184.
- Al Rawashdeh, S. B., 2012, **Assessment of Change Detection Method Based on Normalized Vegetation Index in Environmental Studies**, International Journal of Applied Science and Engineering, Vol. 10(2), pp. 89-97.
- Sabzghabaei, Gh., Dashti, S., Bazm Ara Baleshti, M. & Jafarzadeh K., 2015, **Detecting the variability process of the protected area of Hara Kourkhouran**, Journal of Marine Biology, Vol. 7(26), pp. 1-12.
- Sanhouse-García, A. J., Rangel-Peraza, J. G., Bustos-Terrones, Y., García-Ferrer, A. & Mesas Carrascosa, F. J., 2016, **Land Use Mapping from CBERS-2 Images with Open Source Tools by Applying Different Classification Algorithms**, Physics and Chemistry of the Earth, Parts A/B/C, Vol. 91, pp. 27-37.
- Shao, Y. & Lunetta, R.S., 2012, **Comparison of Support Vector Machine, Neural Network, and CART Algorithms for the Land-Cover Classification Using Limited Training Data Points**, ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 70, pp. 78-87.
- Mantero, P., Moser, G., and Serpico, S. B., 2005, **Partially supervised classification of remote sensing images through SVM-based probability density estimation**, IEEE Transactions on Geoscience and Remote Sensing, Vol. 43(3), pp. 559-570.
- Li, C., Wang, j., Wang, L., Hu, L. & Gong, P., 2014, **Comparison of Classification Algorithms and Training Sample Sizes in Urban Land Classification with Landsat Thematic Mapper Imagery**, Remote Sensing, Vol. 6 (2), pp. 964-983.
- Alizadeh, M., Mirzaie, R. & Kia, H., 2016, **Comparative study of multiple classification methods for land use mapping (Case study: Kann and Karaj rivers basin)**, Geography and Sustainability of Environment, Vol. 20, pp. 89-103.
- Zandieh, V. 2015, **Assessment of land use changes and its impact on groundwater resources based on digital images and GIS analysis in Malayer plain, Hamedan province**, Master thesis, Department of Soil Science at Lorestan University.
- Yousofi, S., Tazeh, M., Mirzaie, S. & Tavangar, Sh., 2014, **Comparison of different satellite image classification algorithms for land use mapping (case study: Noor city)**, Journal of RS and GIS for Natural Resources, Vol. 5(3), pp. 67-76.
- Fathzad, H., Tazeh, M. & Kalantari, S. 2015, **Comparison of the Efficiency of Pixel based Classification Methods (Fuzzy, Arctmap Fuzzy, Neural Networks and Tree Decision) and Object based Methods for Land Use Mapping (Case Study: Dry and Semi-Dry Watershed of Meymeh, Ilam Province)**, Journal of Khoshkboom, Vol. 5(2), pp. 69-82.
- Jamil, A. and Bayram, B. 2018, **Tree Species Extraction and Land Use/Cover Classification from High-Resolution Digital Orthophoto Maps**, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Vol. 11(1), pp. 89-94.

- Kittler, J. and Alkoot, F. M., 2003, **Sum versus vote fusion in multiple classifier systems**, IEEE transactions on pattern analysis and machine intelligence, Vol. 25(1), pp. 110-115.
- James, G. 1998, **Majority vote classifiers: theory and applications** (Doctoral dissertation, Stanford University).
- Abdul Qadir, A., Med-hut, B., & Jirjees, Th., 2010, **Monitoring and evaluation of soil salinity in term of spectral response using landsat images and GIS in Mesopotamian plain/Iraq**, Journal of Iraqi Desert Studies, Vol. 2(2), pp. 19-32.
- Fatemi, B. & Rezaie, Y., 2012, **Fundamental of Remote Sensing. Iran**: Azadeh Press.
- Ren, H., Zhou, G. & Zhang, F., 2018, **Using negative soil adjustment factor in soil-adjusted vegetation index (SAVI) for aboveground living biomass estimation in arid grasslands**, Remote Sensing of Environment, Vol. 209, pp. 439-445.
- Hasanloo, M. & Samadzadegan, F., 2013, **Estimation of inherent dimension in Hyperspectral satellite images**, Journal of Geomatics Science and Technology, Vol. 3(3), pp. 101-109.
- Otukei, J. R. & Blaschke, T., 2010, **Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms**, International Journal of Applied Earth Observation and Geoinformation, Vol. 12, pp. 27-31.
- Mountrakis, G., Im, J., & Ogole, C., 2011, **Support vector machines in remote sensing: A review ISPRS**, Journal of Photogrammetry and Remote Sensing, Vol. 66(3), pp. 247-259.
- Vapnik, W. N., 1999, **An overview of statistical learning theory**, IEEE Transactions of Neural Networks, Vol. 10, pp. 988-999.
- Cortes, C. and Vapnik, V., 1995, **Support-vector networks**, Machine Learning, Vol. 20(3), pp. 273-297.
- Zhu, G. and Blumberg, D. G., 2002, **Classification using ASTER data and SVM algorithms; The case study of Beer Sheva, Israel**, Remote Sensing of Environment, Vol. 80(2), pp. 233-240.
- Mather, P. and Tso, B., 2016, **Classification methods for remotely sensed data**, CRC press.
- Hare, S., Golodetz, S., Saffari, A., Vineet, V., Cheng, M. M., Hicks, S. L. & P. H. Torr, 2016, **Struck: Structured output tracking with kernels**, IEEE transactions on pattern analysis and machine intelligence, Vol. 38(10), pp. 2096-2109.
- Tang, Y., 2013, **Deep learning using linear support vector machines**, arXiv preprint arXiv:1306.0239.
- Zhu, G. B., Liu, X. L. & Jia, Z. G., 2006, **A multi-resolution hierarchy classification study compared with conservative methods**, ISPRS WG II/3, II/6 Workshop Multiple representation and interoperability of spatial data. Hanover, Germany.
- Chasmer, L., Hopkinson, C., Veness, T., Quinton, W., & Baltzer, J., 2014, **A decision-tree classification for low-lying complex land cover types within the zone of discontinuous permafrost**, Remote Sensing of Environment, Vol. 143, pp. 73-84.
- Rouse, Jr. J., Haas, R. H., Schell, J. A., & Deering, D. W., 1974, **Monitoring vegetation systems in the Great Plains with ERTS**.
- Khan, S. I., Hong, Y., Wang, J., Yilmaz, K. K., Gourley, J. J., Adler, R. F., & Irwin, D., 2011, **Satellite remote sensing and hydrologic modeling for flood inundation mapping in Lake Victoria basin: Implications for hydrologic prediction in ungauged basins**, IEEE Transactions on Geoscience and Remote Sensing, Vol. 49(1), pp. 85-95.
- Mather, M. M., 1999, **Computer processing of remotely sensed data**, 2<sup>nd</sup> Edition, John Wiley & Sons Press.
- Arkhi, S., 2012, **Assessment of the Effectiveness of Decision Tree Classification Method for Extracting Land uses Map by Using Satellite Data in Cham Gardalan Catchment Area**, Geography and Territorial Spatial Arrangement, Vol. 4, pp. 17-26.





## A hybrid classification method based on fusion of parametric and non-parametric classification algorithms for Landuse/Landcover map in Hirkani Forests

Saadat, M.<sup>1</sup> and Shah-Hosseini, R.\*<sup>2</sup>

1. MSc Student in Remote Sensing Engineering, School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran
2. Assistant Professor, School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran

### Abstract

Preparation of proper land use maps has always been one of the important goals of researchers and policymakers. The aim of this study was to provide a new method for preparing land use maps using remotely sensed data and satellite data imagery. For this Purpose, we used Landsat 8 data, Digital Elevation Model (DEM), Principal Component Analysis (PCA), and Spectral Indices to extract land use map in the study area. After all required preprocessing, the training samples were provided. In this study, the training samples were utilized in two parts; in the first part they were used as inputs for image classification using supervised algorithms of maximum likelihood Classification (MLC) and support vector machine (SVM). In the second part, in order to applying Decision Tree Classification (DTC), these training samples were used to determine the spectral reflection of each end-member in the spectrum of electromagnetic waves (image bands, PCA, spectral indices, and DEM). Then, using these binary data and DTC, each end-member was identified and the Landuse/Landcover (LULC) map was extracted. In order to combine the classification results and achieve higher accuracy, the Majority Vote Classification (MVC) method was applied to prepare a new compilation of land use in the area. In order to evaluate the accuracy of produced maps, the statistical parameters extracted from the confusion matrix including overall accuracy, kappa coefficient, user and producer's accuracy were utilized. According to the results, the combined method (MVC) with a total accuracy of 93.37% and kappa coefficient of 0.91 had the highest accuracy. The overall accuracy of the DTC, SVM, and MLC were 89.61, 88.01 and 87.6%, respectively. Due to the fact that in the nature most of the landuse are mixed and complicated, it would be better to use new methods that cover all aspects of the phenomena. In this research, the data extracted from the supervised classifications as well as the data derived from the DTC were combined and the results clearly illustrate the improvement of the final accuracy of the classification.

**Keywords:** Majority Vote Classification, Landuse, Classification Algorithm, Remote Sensing, Landsat8