بررسی عملکرد روش‌های یادگیری جمعی با توجه به روش انتخاب ویژگی، به منظور ادغام طبقه‌بندی کننده‌های انعکاسی و حرارتی با هدف شناسایی ابر، ابر سیروس و برف/یخ در تصاویر مادیس

نفیسه قاسمیان*[۱]، مهدی آخوندزاده هنزائی[۲]

[۱] دانشجوی کارشناسی ارشد سنجش از دور - دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی - دانشگاه تهران
n.ghasemian@ut.ac.ir

[۲] استادیار دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی - دانشگاه تهران
makhonz@ut.ac.ir

(تاریخ دریافت آبان ۱۳۹۵، تاریخ تصویب فروردین ۱۳۹۶)

چکیده

تقریبا همه‌ی تصاویر سنجنده‌ی مادیس دارای قسمت‌های پوشیده از ابر هستند. به منظور استخراج اطلاعات صحیح از داده‌های مادیس، یکی از پیش‌پردازش‌های کلیدی شناسایی پیکسل‌های ابری و جداسازی آن از عوارض مشابه مانند برف/یخ است. ویژگی‌های مورد استفاده در طبقه-بندی ابر به دو دسته‌ی ویژگی‌های بافتی و طیفی تقسیم می‌شوند. با استفاده از ویژگی‌های بافتی باندهای مرئی امکان جداسازی پیکسل‌های ابر از پیکسل‌های برف/یخ فراهم می‌شود ولی ابر و برف می‌توانند دارای ویژگی‌های حرارتی مشابه باشند. همچنین از ویژگی‌های حرارتی (دما) در ماسک ابر مادیس به منظور شناسایی ابرها در ارتفاع‌های مختلف استفاده شده است. مطالعات زیادی به منظور طبقه‌بندی پوشش سطح زمین با استفاده از روش‌های یادگیری جمعی انجام شده است و از این روش‌ها صرفا به منظور طبقه‌بندی استفاده شده است. در این تحقیق کاربردی جدید از روش‌های یادگیری جمعی در مقایسه با مطالعات پیشین مطرح شده است و از این روش‌ها به منظور ادغام دو نوع مختلف از طبقه‌بندی کننده‌ها که نوع اول طبقه‌بندی کننده‌هایی با ویژگی‌های انعکاسی و نوع دوم با ویژگی‌های حرارتی هستند، استفاده شده است. همچنین در مطالعات پیشین، اثر تغییر ویژگی‌های ورودی بر عملکرد نهایی روش‌های یادگیری جمعی مورد بررسی قرار نگرفته است. بنابراین هدف این تحقیق مقایسه-ی نتیجه‌ی ادغام طبقه‌بندی کننده‌های با ویژگی‌های انعکاسی و حرارتی با استفاده از دو نوع از روش‌های یادگیری جمعی شامل boosting و الگوریتم جنگل تصادفی(RF)، به منظور شناسایی پیکسل‌های ابری، سیروس و برف/یخ با توجه به روش انتخاب ویژگی می‌باشد. ابتدا به منظور انتخاب ویژگی‌های انعکاسی و حرارتی در روش‌های boosting به کار گرفته شده، شامل adaboost.M1، adaboostSVM، logitboost و totalboost از روش‌های معیار S و الگوریتم ژنتیک (GA) و در روش RF علاوه بر روش‌های ذکر شده از روش حذف ویژگی به روش بازگشتی (RFE) و ماتریس کارلیشن استفاده شد. سپس طبقه‌بندی کننده‌ها در سطح تصمیم با یکدیگر ادغام شدند. برای اکثر روش‌های یادگیری جمعی صرف نظر از روش انتخاب ویژگی، دقت تولیدکننده‌ی ابر وسیروس بالایی دست آمد. استفاده از دو روش RFE و ماتریس کارلیشن در الگوریتم RF توانست دقت کاربری پیکسل‌های ابر به ترتیب ۹۹٪ و ۱۰۰٪ را نتیجه دهد که نسبت به حالتی که از روش‌های معیار S و الگوریتم ژنتیک (GA) برای انتخاب ویژگی استفاده شد، دقت‌های بالاتری را نشان داد. روش‌های boosting صرف نظر از روش انتخاب ویژگی با اختصاص وزن بیشتر به داده‌های آموزشی مربوط به کلاس با تعداد داده‌های آموزشی کم‌تر، توانستند به دقت تولیدکننده‌ی برف/یخ بالاتری تسبت به الگوریتم RF دست یابند. همچنین این روش‌ها دقت کاربری سیروس نسبتا بالاتری نسبت به روش‌های RF نتیجه دادند. در بین روش‌های انتخاب ویژگی مختلف در RF روش ماتریس کارلیشن توانست دقت کاربری سیروس ۹۱٪ را نتیجه دهد. در انتها، میزان توافق نتایج طبقه‌بندی با نقشه‌ی مرجع به دست آمده از ماسک ابر مادیس محاسبه شد. روش‌های RF درصد توافق‌های بالاتری نسبت به روش‌های boosting نتیجه دادند. بالاترین درصد توافق برای روش RF-RFE به مقدار ۷۶٪ و پایین‌ترین برای روش logit boost-GA به مقدار ۴۲٪ به دست آمد.

واژگان کلیدی: روش‌های یادگیری جمعی، انتخاب ویژگی، ابر، برف/یخ، سیروس، ادغام، طبقه‌بندی کننده‌های انعکاسی و حرارتی

* نویسنده رابط

۱- مقدمه

اولین گام در طبقه‌بندی پیدا کردن ویژگی‌های مناسب است. بسیاری از ویژگی‌های ورودی به فرایند طبقه‌بندی برای شناسایی عارضه‌ی مورد نظر مناسب نیستند. ویژگی نامناسب به ویژگی گفته می‌شود که استفاده از آن در طبقه‌بندی و شناسایی عارضه‌ی مورد نظر تاثیری نداشته باشد و یا اطلاعات جدیدی در مورد تارگت اضافه نکند[۱]. در بسیاری از مواقع حجم داده‌های ورودی به طبقه‌بندی بالا بوده و فرایند یادگیری قبل از حذف ویژگی‌های اضافی به خوبی انجام نمی‌شود. حذف این ویژگی‌های نامرتبط و اضافی، موجب افزایش سرعت یادگیری می‌شود. بنابراین روش‌های انتخاب ویژگی سعی می‌کنند مجموعه‌ای از ویژگی‌های مرتبط با شناسایی عوارض مورد نظر را انتخاب کنند. تنوع روش‌های انتخاب ویژگی موجود این سوال را به وجود می‌آورد که کدام یک از روش‌های انتخاب ویژگی متداول، در کاربرد مورد نظر مناسب است. بر اساس نحوه‌ی تولید ویژگی‌ها و تابع ارزیابی می‌توان روش‌های متنوعی برای انتخاب ویژگی در نظر گرفت. روش‌های مختلفی برای تولید زیر مجموعه‌ای از ویژگی‌ها وجود دارند که عبارتند از:۱- جستجوی کامل ۲- جستجوی اکتشافی[۱] ۳- جستجوی رندم. روش جستجوی رندم، روشی نسبتا جدید نسبت به دو روش دیگر می‌باشد. هرچند ابعاد فضای جستجو 2^N است ولی در عمل تعداد ویژگی‌هایی که مورد بررسی قرار می‌گیرند کم‌تر از 2^N می‌باشد و الگوریتم پس از تعداد تکرار معین متوقف می‌شود. عملکرد این روش بستگی به ویژگی‌های ورودی و نحوه‌ی تولید اعداد تصادفی دارد. از این دسته می‌توان به روش‌های تصادفی مانند الگوریتم ژنتیک اشاره کرد. پیدا کردن یک زیر مجموعه‌ی بهینه با توجه به یک تابع ارزیاب انجام می‌شود. تابع ارزیابی، قابلیت جداسازی کلاس‌های مختلف، برای یک ویژگی یا مجموعه‌ای از ویژگی‌ها را اندازگیری می‌کند. تابع‌های ارزیاب را می‌توان به پنج دسته تقسیم کرد[۱]: تابع فاصله، اطلاعات، وابستگی، هماهنگی[۲] و میزان خطای طبقه‌بندی.

روش‌های شناسایی ابر را می‌توان به دو دسته‌ی روش‌های فیزیکی و مدل‌های داده کاوی[۳] تقسیم کرد[۲]. از روش‌های دسته‌ی اول می‌توان به روش‌های ذکر شده در منابع [۳]-[۸] اشاره کرد. در سال‌های اخیر تحقیقات قابل توجهی در حیطه‌ی شناسایی ابر با استفاده از روش‌های داده کاوی انجام شده است. در مقالات از طبقه‌بندی کننده[۴]هایی مانند شبکه عصبی[۵]، احتمال بیشینه[۶]، درخت تصمیم‌گیری[۷] و نزدیک‌ترین همسایگی[۸] برای شناسایی ابر استفاده شده است. در مقایسه با روش‌های فیزیکی، روش‌های داده کاوی دقت طبقه‌بندی را به طور قابل توجهی بهبود داده‌اند. همچنین تحقیقات زیادی به منظور شناسایی ابر در تصاویر سنجنده‌ی مادیس انجام شده است[۹].

ویژگی‌های مورد استفاده در طبقه‌بندی ابر به دو دسته‌ی ویژگی‌های طیفی و بافتی تقسیم می‌شوند. دسته‌ی اول که نقش مهم‌تری در طبقه‌بندی ابر ایفا می‌کنند، اطلاعات مرتبط با رادیانس ابر را در باندهای طیفی مختلف استخراج می‌کند[۱۰]. به عنوان مثال از این دسته روش‌ها می‌توان به روش‌های مبتنی بر حدآستانه، روش‌های هیستوگرام مبنا و روش‌های چند طیفی اشاره کرد. ویژگی‌های طیفی به خاطر اهمیت فیزیکی‌شان (انعکاس و دما) ویژگی‌های ساده و در عین حال موثری برای شناسایی ابر می‌باشند. هر چند این ویژگی‌ها در تشخیص ابرهای یخی[۹] و برف به خاطر شباهت طیفی‌شان با مشکل مواجه می‌شوند. دسته‌ی دوم، انواع مختلف ابر را با استفاده از توزیع مکانی وقوع درجات خاکستری در یک ناحیه و در یک باند خاص، شناسایی می‌کند. از جمله ویژگی‌های بافتی مفید در شناسایی ابر می‌توان به دو ویژگی کنتراست و وابستگی اشاره کرد[۱۱]. در مطالعه‌ای صحت طبقه‌بندی ابر در تصاویر ماهواره‌ی لندست، با دو روش مختلف در استخراج بافت، با یکدیگر مقایسه شدند. نتایج به دست آمده صحت پایین‌تری را برای روش GLDV[۱۰] نسبت به روش GLCM[۱۱] نشان داد[۱۲]. چانگ هو آی و همکارانش از ویژگی‌های بافت

۱ heuristic
۲ consistency

۳ Data mining models
٤ classifier
٥ Neural network
٦ Maximum likelihood
۷ Decision tree
۸ 1-nearest neighbors
۹ Ice clouds
۱۰ GLDV
۱۱ Gray level co-occurrence matrix

ماتریس GGCM[1] به منظور شناسایی ابر استفاده کردند[۱۳]. از معایب این روش می‌توان به در نظر گرفتن عوارضی که ویژگی‌های مشابه با ابر دارند، مانند: بیابان، مه، برف و دود، به عنوان ابر اشاره کرد.

در سال‌های اخیر تحقیقات زیادی در زمینه‌ی ترکیب طبقه‌بندی کننده‌ها به منظور ایجاد یک طبقه‌بندی کننده‌ی نهایی پیشنهاد داده شده‌است. این طبقه‌بندی کننده معمولا از طبقه‌بندی کننده‌های اولیه صحت بالاتری دارد. به روش‌های ترکیب طبقه‌بندی کننده‌ها روش‌های یادگیری جمعی[2] گفته می‌شود. در مطالعه‌ای از روش‌های یادگیری جمعی bagging و boosting به منظور طبقه‌بندی پوشش زمین با استفاده از طبقه‌بندی کننده‌ی ماشین بردار پشتیبان (SVM) استفاده شد. نتایج به دست آمده نشان داد که روش boosting عملکرد طبقه‌بندی کننده‌ی SVM را بهبود نداد[۱۴]. همچنین در مطالعه‌ی دیگری عملکرد سه روش یادگیری جمعی، bagging، boosting و الگوریتم جنگل تصادفی (RF) در طبقه‌بندی پوشش اراضی، با روش درخت تصمیم‌گیری مقایسه شد. نتایج به دست آمده نشان داد که مقادیر ضریب کاپای به دست آمده با استفاده از روش‌های یادگیری جمعی ۱۱٪ بیش از مقدار مشابه به دست آمده با استفاده از الگوریتم درخت تصمیم‌گیری بود[۱۵]. در تحقیقات ذکر شده، اثر تغییر روش انتخاب ویژگی بر عملکرد نهایی روش‌های یادگیری جمعی مورد بررسی قرار نگرفته است. همچنین در این مطالعات از روش‌های یادگیری جمعی صرفا به منظور طبقه‌بندی پوشش سطح زمین استفاده شده و کارایی این روش‌ها در ادغام طبقه‌بندی کننده‌ها بررسی نشده‌است. مطالعات انجام شده در طبقه‌بندی با استفاده از مجموعه‌ای از طبقه‌بندی کننده‌ها و یا تنها یک طبقه‌بندی کننده، وابستگی نتایج را به تعداد و نوع ویژگی‌های ورودی نشان می‌دهند[۱۶]–[۱۸]. بنابراین با توجه به تحقیقات انجام شده می‌توان گفت روش انتخاب ویژگی نیز می‌تواند یک پارامتر تاثیر گذار در نتیجه‌ی طبقه‌بندی سیستم‌های طبقه‌بندی کننده‌ی چندگانه[3] باشد.

در این مقاله عملکرد ۴ نوع مختلف از روش‌های boosting شامل، adaboost.M1، adaboostSVM، logitboost، totalboost و الگوریتم جنگل تصادفی به منظور ادغام طبقه‌بندی کننده‌های انعکاسی (ویژگی‌های بافتی و طیفی انعکاس باندهای مرئی-مادون قرمز نزدیک) و حرارتی (ویژگی‌های بافتی و طیفی مقادیر دما) با توجه به روش انتخاب ویژگی،یا هدف شناسایی پیکسل‌های ابر، برف/یخ و سیروس با هم مقایسه شدند. همچنین میزان توافق نتایج طبقه‌بندی با روش‌های مختلف انتخاب ویژگی، با ماسک ابر مادیس، محاسبه شد.

۲- مبانی تئوری تحقیق

به منظور شناسایی پیکسل‌های ابر، برف/یخ و سیروس از روش‌های adaboost.M1، adaboostSVM، logitboost، $totalboost_v$ و الگوریتم جنگل تصادفی (RF) استفاده شد، بنابراین بررسی اجمالی مبانی تئوری این روش‌ها ضروری به نظر می‌رسد.

۲-۱-Adaboost.M1

Adaboost.M1 رایج‌ترین نوع از روش‌های آدابوست می‌باشد. جزئیات این الگوریتم در مرجع[۱۹] بیان شده است.

۲-۲- AdaboostSVM

طبقه‌بندی کننده‌ی SVM یک طبقه‌بندی کننده‌ی پارامتریک است. کرنل گوسین شامل دو پارامتر پهنای گوسین σ و ضریب تنظیم‌گر C می‌باشد. تغییر در مقادیر این پارامترها موجب تغییر در عملکرد طبقه‌بندی کننده می‌شود. هر چند انتخاب مقداری کوچک برای پارامتر تنظیم‌گر موجب کاهش دقت عملکرد طبقه‌بندی با کرنل گوسین می‌شود ولی بررسی‌ها نشان می‌دهد که در صورت انتخاب یک مقدار مناسب برای پارامتر تنظیم‌گر، عملکرد کرنل گوسین بیشتر تحت تاثیر پارامتر σ می‌باشد[۲۰]. مقدار خطای تست نهایی وابستگی زیادی به مقدار C ندارد. همچنین پارامتر σ_{step} تعداد دفعات تکرار الگوریتم را تعیین می‌کند و و تغییر آن تاثیر چندانی بر روی مقدار خطای تست نهایی ندارد. برای توضیحات بیش‌تر درباره این روش به مرجع [۲۰] مراجعه شود.

[1] Gradient gray level co-occurrence matrix
[2] ensemble learning methods
[3] Multiple classifier systems (MCS)

Logitboost -۳-۲

الگوریتم‌های آدابوست می‌توانند به عنوان فرآیندهای گام به گام به منظور برازش مدل لجستیک افزایشی در نظر گرفته شوند[۲۱]. این الگوریتم یک معیار نمایی را بهینه می‌کند. بسط سری تیلور این معیار تا مرتبه‌ی دوم برابر با معیار احتمال لگاریتمی دو جمله‌ای است. هدف در این روش، پیدا کردن تابع F(x) ای است که رابطه‌ی (۱) را مینیمم کند.

$$J(F) = E(e^{-yF(x)}) \qquad (۱)$$

در رابطه‌ی(۱) E می‌تواند امید ریاضی جامعه‌ی آماری با توزیع احتمال مشخص و یا میانگین نمونه باشد. تابع F(x)ای که مقدار رابطه‌ی (۱)را مینیمم می‌کند، تبدیل لجستیک P(y==1|x) است (رابطه‌ی (۲)).

$$F(x) = \frac{1}{2} log \frac{P(y=1|x)}{P(y=-1|x)} \qquad (۲)$$

بنابراین

$$P(y=1|x) = \frac{e^{F(x)}}{e^{-F(x)} + e^{F(x)}} \qquad (۳)$$

$$P(y=-1|x) = \frac{e^{-F(x)}}{e^{-F(x)} + e^{F(x)}} \qquad (۴)$$

totalboost$_v$ -۴-۲

روش بوستینگ می‌تواند به عنوان یک مسئله‌ی بهینه سازی در نظر گرفته شود. هدف این الگوریتم انتخاب یک ترکیب کوژ[۱] از طبقه‌بندی کننده‌های به اصطلاح ضعیف است، به گونه‌ای که مقدار حاشیه حداکثر شود. الگوریتم آدابوست که یکی از انواع رایج روش بوستینگ می‌باشد، می‌تواند به عنوان مسئله‌ای با هدف مینیمم کردن آنتروپی توزیع وزن داده‌های آموزشی در آخرین تکرار نسبت به توزیع وزن دراولین تکرار در نظر گرفته شود. این روند مینیمم سازی تا زمانی ادامه می‌یابد که میانگین وزن‌دار حاشیه‌ی نمونه‌ها (edge) برابر با مقدار صفر شود (این شرط معادل با نصف شدن خطای وزن دار طبقه‌بندی کننده در انواع پیشین بوستینگ می‌باشد). در این روش فرآیند مینیمم سازی تابع آنتروپی نسبی با این شرط انجام می‌شود که مقادیر edge در t طبقه‌بندی گذشته، از مقدار γ بیشتر نشود. به چنین الگوریتم‌هایی در اصطلاح "totally corrective" گفته می‌شود[۲۲]. این روش‌ها به تعداد دفعات تکرار کم‌تری نیاز دارند و روش‌های مناسبی به منظور انتخاب ویژگی[۲] می‌باشند.

۵-۲- الگوریتم جنگل تصادفی[۳]

الگوریتم جنگل تصادفی از جدیدترین روش‌های یادگیری جمعی است[۲۳]. این روش تعمیم الگوریتم bagging می‌باشد و تفاوت اصلی آن در انتخاب ویژگی تصادفی است. در هر گره[۴] مجموعه‌ای از ویژگی‌ها به طور تصادفی انتخاب می‌شوند و سپس مراحل انتخاب ویژگی در هر گره، در فضای ویژگی ادامه می‌یابد. پارامتر K در این الگوریتم میزان مشارکت تصادفی بودن را کنترل می‌کند. در صورتی که K=1 در نظر گرفته شود، در هر گره یک ویژگی انتخاب می‌شود؛ در صورتی که K برابر با کل ویژگی‌های موجود در نظر گرفته شود، درخت تصمیم ایجاد شده همان درخت تصمیم قطعی قدیمی خواهد بود. مقدار پیشنهادی برای K برابر با لگاریتم تعداد ویژگی‌ها می‌باشد، همچنین در برخی منابع مقدار این پارامتر برابر با جذر تعداد ویژگی‌ها یا باندهای طیفی در نظر گرفته شده‌است[۲۴]. تصادفی بودن این روش به مرحله‌ی انتخاب ویژگی مربوط می‌شود و انتخاب حدآستانه در هر گره یک پروسه‌ی تصادفی نمی‌باشد[۲۵].

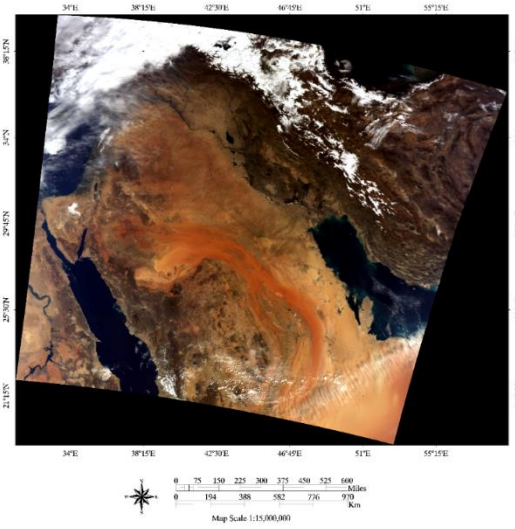
۳- داده‌های مورد استفاده و منطقه‌ی مورد مطالعه

به منظور شناسایی پیکسل‌های ابری از داده‌های MOD021KM-L1B ماهواره‌ی Terra سنجنده‌ی مادیس استفاده شد. این داده‌ها دارای رزولوشن مکانی KM ۱ هستند و باید قبل از استفاده، بر روی‌آن‌ها به ترتیب تصحیحات

۱ y(v,w,x) یک ترکیب کوژاز نقاط a,b,c نامیده می‌شود در صورتی که $a,b,c \in \mathbb{R}^*$ and $v,w,x \in \mathbb{R}, y(v,w,x) = va + wb + xc,\ v+w+x=1,\ v,w,x \geq 0$

۲ Feature selection
۳ Random forest
۴ node

هندسی و رادیومتریکی صورت گیرد. همچنین تصاویر باید زمین مرجع شوند. تصاویر مادیس دارای ۳۶ باند طیفی در محدوده‌ی µm ۱۴-۴/. می‌باشند. تصویر مورد استفاده مربوط به منطقه‌ی خاورمیانه بوده و در تاریخ سوم فوریه‌ی سال ۲۰۱۶ ساعت ۷:۵۰ صبح اخذ شده است. شکل ۱ محدوده‌ی منطقه‌ی مورد مطالعه را بر روی نقشه نشان می‌دهد. در این تحقیق به منظور ارزیابی صحت پیکسل‌های ابری شناسایی شده نتایج با نقشه‌ی مرجع به دست آمده از MOD35 مقایسه شدند. بنابراین در این تحقیق از داده‌های MOD35 نیز استفاده شد. داده‌های مورد نیاز از سایت NASA به آدرس اینترنتی http:www.ladsweb.nascom.nasa.gov به صورت رایگان در فرمت hdf قابل دانلود می‌باشند.



شکل ۱- محدوده‌ی منطقه‌ی مورد مطالعه

۴- پیش پردازش

همان طور که گفته شد، قبل از انجام پردازش‌های لازم بر روی داده‌های Level 1B سنجنده‌ی مادیس، نیاز به انجام تصحیحات هندسی، رادیومتریکی و تبدیل مقادیر DN به مقادیر انعکاس و دما می‌باشد. تصحیح هندسی تصاویر مادیس شامل حذف خطای bow tie می‌شود. جهت تصحیح خطای bow tie از افزونه‌ی Modis conversion toolkit(MCTK) که بر روی نرم‌افزار ENVI نصب می‌شود، استفاده شد. همزمان با حذف خطای bow tie سیستم مختصات تصاویر از سیستم مختصات محلی به جهانی انتقال یافت. تصحیح رادیومتریکی به معنی کاهش اثر خطای نوار نوار شدگی[۱] می‌باشد[۲۶]-[۲۷].

در این تحقیق، از روش تطبیق هیستوگرام[۲] به منظور کاهش خطای نوار نوار شدگی استفاده شد[۲۸]. سپس مقادیر DN با استفاده از نرم افزار ENVI به انعکاس بالای اتمسفر و دمای درخشندگی تبدیل شدند. برای اختصار در این مقاله، به جای واژه‌های انعکاس بالای اتمسفر و دمای درخشندگی به ترتیب از انعکاس و دما استفاده شده است.

۵- روش تحقیق

بعد از انجام پیش‌پردازش‌های لازم و کاهش خطاهای ذکر شده، ویژگی‌های بافتی و طیفی از تصویر استخراج شدند. ویژگی‌های بافتی استفاده شده در این تحقیق عبارتند از همه‌ی ویژگی‌های بافتی انعکاسی استخراج شده از ماتریس رخداد توام در همسایگی‌های ۳×۳ شامل میانگین، وریانس، همگنی، کنتراست، عدم شباهت، آنتروپی، گشتاور مرتبه دوم و وابستگی و در مورد باندهای حرارتی از برخی از ویژگی‌های بافت شامل وابستگی، میانگین،گشتاور مرتبه دوم و عدم شباهت استفاده شد. ویژگی‌های طیفی به کار گرفته شده شامل مقادیر انعکاس، دمای درخشندگی و اختلافات مقادیر دمای درخشندگی می‌باشند. جدول۱ لیست ویژگی‌های انعکاسی و حرارتی ورودی را نشان می‌دهد. باندهای ۱و۲ از باندهای مفید در شناسایی ابر می‌باشند. همچنین باند ۸ در محدوده‌ی طول موج آبی بوده و به منظور بررسی قابلیت باندهای مرئی در شناسایی ابر در نظر گرفته شد به علاوه این باند دارای کاربرد برای شناسایی ابر در مناطق بیابانی است[۲۹]. باند ۲۶ نیز در ماسک ابر مادیس به منظور شناسایی ابر سیروس مورداستفاده قرار گرفته‌است[۳۰]. همچنین شاخص‌های NDVI، NDSI و نسبت انعکاس باند ۱ به باند ۲ به ترتیب دارای کاربرد در شناسایی ابر در مناطق دارای پوشش گیاهی، برفی و مناطق بیابانی و آبی می‌باشند. ویژگی‌های حرارتی در نظر گرفته شده به عنوان ورودی عبارتند از: دمای باند۳۱،۲۲،۲۰ و ۳۲ به دلیل کاربرد در شناسایی دمای ابر، ویژگی‌های بافت این باندها و اختلاف‌های دمای درخشندگی شامل۱- BT_{31}-BT_{32} ۲- BT_{33}-BT_{31} ۳- BT_{31}-BT_{20} ۴- BT_{20}-BT_{32} ۵- BT_{22}-BT_{32} ۶- دمای باند ۳۵. ویژگی‌های ۵،۴،۲،۱ و ۶ به منظور شناسایی ابرهای ارتفاع بالا و ویژگی۳، برای شناسایی ابرهای ارتفاع بالا و پایین در ماسک ابر مادیس استفاده شده است[۳۰].

۱ stripping

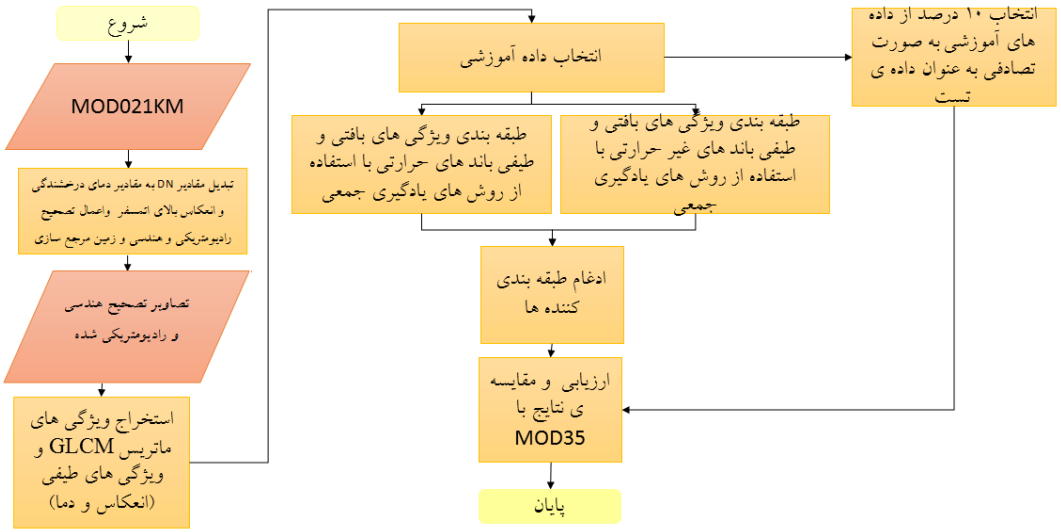
۲ Histogram matching

جدول۱- ویژگی‌های انعکاسی و حرارتی ورودی

ویژگی‌های انعکاسی		ویژگی‌های حرارتی	
ویژگی‌های بافت	ویژگی‌های طیفی	ویژگی‌های بافت	ویژگی‌های طیفی
ویژگی‌های بافت ماتریس رخداد توام باندهای ۱،۲،۸ و ۲۶ (۸ ویژگی برای هر باند)	انعکاس باندهای ۱، ۲، ۸ و باند ۲۶	کارلیشن باند ۲۰، میانگین و گشتاور مرتبه دوم باند ۲۲، میانگین و عدم شباهت باند ۳۱ و میانگین باند ۳۲	دمای باندهای ۲۰،۲۲،۳۱،۳۲ و ۳۵
	شاخص‌های NDVI، NDSI و نسبت انعکاس باند ۱ به ۲		$BT_{۳۱}$-$BT_{۳۲}$ $BT_{۳۳}$-$BT_{۳۱}$ $BT_{۳۱}$-$BT_{۲۰}$ $BT_{۳۲}$-$BT_{۲۰}$ $BT_{۲۲}$-$BT_{۳۲}$

سپس ویژگی‌های بافتی و طیفی انعکاسی و حرارتی ذکر شده، به طور جداگانه وارد فرایند انتخاب ویژگی شدند. به منظور انتخاب ویژگی در طبقه‌بندی با انواع مختلف boosting از معیار جدایی پذیری (S) و الگوریتم ژنتیک و در طبقه‌بندی با استفاده از الگوریتم جنگل تصادفی علاوه بر روش‌های فوق، ماتریس کارلیشن و حذف ویژگی به صورت بازگشتی (موجود در بسته‌ی caret نرم‌افزار R) نیز استفاده شدند. سپس بر روی ویژگی‌های استخراج شده از باندهای غیر حرارتی با استفاده از ۴ نوع مختلف از الگوریتم boosting شامل adaboost.M1، adaboostSVM، logitboost، totalboost و الگوریتم جنگل تصادفی طبقه‌بندی انجام شد. بر روی ویژگی‌های استخراجی از باندهای حرارتی نیز با استفاده از روش‌های ذکر شده طبقه‌بندی انجام و طبقه‌بندی کننده‌ها در سطح تصمیم با توجه به نوع خروجی با استفاده از دو روش رای اکثریت و احتمال بیشینه (در روش logitboost)، با یکدیگر ادغام شدند. روش‌های boosting در محیط نرم-افزار MATLAB و الگوریتم جنگل تصادفی در محیط نرم‌افزار R پیاده سازی شدند. در انتها از ماسک ابر مادیس(Mod35) به عنوان رفرنس استفاده شده و نتایج با ماسک ابر مادیس مقایسه شد (شکل ۲).

شکل۲- فلوچارت مراحل انجام تحقیق



۵-۱- انتخاب ویژگی با استفاده از معیار S

می‌توان از معیار جدایی پذیری که بر اساس میانگین و انحراف معیار ویژگی‌ها بین دو کلاس محاسبه می‌شود (رابطه‌ی 5)، به منظور انتخاب ویژگی استفاده کرد. در رابطه‌ی (۵) S_{ijk} میزان جدایی‌پذیری ویژگی v_i بین کلاس‌های C_j و C_k می‌باشد. $\overline{m_{i,j}}$ و $\sigma_{i,j}$ به ترتیب میانگین و انحراف معیار ویژگی i در کلاس Cj هستند. هر چه توصیف‌گر در نظر گرفته شده مقدار Sepratability بالاتری داشته باشد، بدین معنی است که دو کلاس مورد نظر را بهتر از سایر ویژگی‌ها از یکدیگر تشخیص می‌دهد.

با محاسبه‌ی مقدار پارامتر S بین هر کدام از زوج کلاس‌ها و برای هر کدام از ویژگی‌ها به طور جداگانه می‌توان مجموعه توصیف‌گرهایی که دارای بیشترین مقدار پارامتر S بین هر کدام از زوج کلاس‌ها هستند را به عنوان ویژگی‌های بهینه انتخاب کرد. به عنوان مثال مقادیر پارامتر S برای دو ویژگی NDVI و NDSI بین هر کدام از زوج کلاس‌ها در جداول ۲ و ۳ آمده‌است.

$$S_{i,j,k} = \frac{\left|\overline{m_{i,j}} - \overline{m_{i,k}}\right|}{\sigma_{i,j} + \sigma_{i,k}} \qquad (۵)$$

به دلیل یکسان بودن پارامتر S بین کلاس‌های j وk وکلاس kو j ماتریس‌های نشان داده شده در جداول ۲ و ۳ ماتریس‌هایی متقارن هستند بنابراین تنها درایه‌های بالای قطر اصلی نشان داده شده‌اند. به عنوان مثال با توجه به جداول ۲و ۳ ملاحظه می‌شود که میزان پارامتر S بین کلاس‌های ۱و ۶ (کلاس ابر و ابر سیروس) برای شاخص NDSI مقدار ۲/۳۷ و برای شاخص NDVI برابر با ۰/۸۲ می‌باشد بنابراین شاخص NDSI در تشخیص این دو کلاس بهتر عمل می‌کند. با مقایسه‌ی پارامتر S بین دو کلاس ۱و ۶ برای سایر ویژگی‌ها، در نهایت ویژگی که دارای بیشترین مقدار S است، انتخاب می‌شود. این کار برای سایر جفت کلاس‌ها تکرار شده و در نهایت مجموعه‌ای از ویژگی‌ها که دارای ماکزیمم مقدار پارامتر S بین جفت کلاس‌ها هستند انتخاب می‌شوند. جدول ۶ ویژگی‌های انعکاسی و حرارتی انتخاب شده با استفاده از روش‌های معیار S و سایر روش‌های به کار گرفته شده شامل الگوریتم ژنتیک، ماتریس کارلیشن و روش بازگشتی را نشان می‌دهد. با توجه به این جدول ملاحظه می‌شود که در چهار روش انتخاب ویژگی ذکر شده خروجی‌های متفاوتی به عنوان ویژگی‌های بهینه به دست آمده است. علت این اختلاف تفاوت در تابع هدف می‌باشد. برای مثال تابع هدف در روش معیار S، معیار جدایی پذیری است که از میانگین و انحراف معیار ویژگی‌ها برای تعریف آن استفاده می‌شود و یا در الگوریتم ژنتیک تابع هدف میزان عدم توافق نتیجه‌ی طبقه‌بندی SVM با نقشه‌ی رفرنس به دست آمده از ماسک ابر مادیس در نظر گرفته شد و یا در روش بازگشتی با استفاده از معیار اهمیت ویژگی به هر مجموعه از ویژگی‌ها رنک داده می‌شود. بنابراین این تفاوت در تابع هدف منجر به انتخاب ویژگی‌های متفاوت و در نتیجه تفاوت در دقت شناسایی کلاس‌های هدف می‌شود. این مطلب می‌تواند در شناسایی عارضه‌ی مورد نظر کمک کننده باشد.

جدول۲- معیار جدایی پذیری (S) برای ویژگی NDVI

NDVI	SNDVI,۱	SNDVI,۲	SNDVI,۳	SNDVI,۴	SNDVI,۵	SNDVI,۶	SNDVI,۷
SNDVI,۱	۰	۰/۱۹۲۳	۰/۰۳۲۳	۱/۴۶۸۴	۰/۲۳۹۷	۰/۸۱۹۷	۲/۳۹۳۳
SNDVI,۲			۰/۰۶۴۹	۲/۲۰۱۹	۱/۰۵۲۷	۱/۱۹۷۹	۷/۹۴۵۷
SNDVI,۳				۱/۰۳۴۵	۰/۱۶۹	۱/۴۲۷۴	۱/۱۹۷۲
SNDVI,۴					۲/۰۷۲۹	۲/۵۹۳۷	۰/۷۸۳۶
SNDVI,۵						۲/۳۸۲۹	۱۱/۸۳۸۱
SNDVI,۶							۸/۲۳۶۶
SNDVI,۷							۰

جدول۳- معیار جدایی پذیری (S) برای ویژگی NDSI

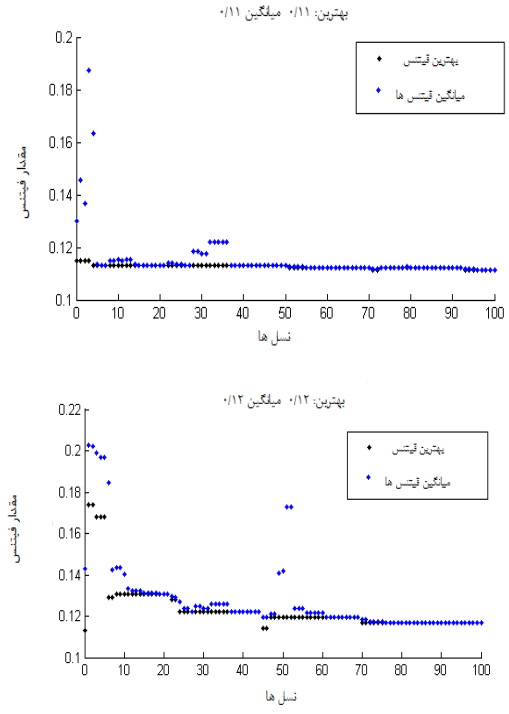
NDSI	SNDSI,۱	SNDSI,۲	SNDSI,۳	SNDSI,۴	SNDSI,۵	SNDSI,۶	SNDSI,۷
SNDSI,۱	۰	۱/۷۹۶	۰/۸۳۱۱	۰/۵۳۴۸	۲/۹۶۹	۲/۳۶۶۶	۰/۶۰۳۷
SNDSI,۲			۱/۹۹۱۴	۰/۷۸۳۳	۷/۳۹۵۸	۵/۷۹۴۹	۳/۹۷۰۹
SNDSI,۳				۱/۱۰۰۹	۰/۵۷۵۲	۰/۳۹۷	۱/۴۹۱۷
SNDSI,۴					۳/۰۵۳۹	۲/۵۵۹۸	۰/۴۸۴۸
SNDSI,۵						۰/۳۹۵۹	۱۲/۲۹۹۱
SNDSI,۶							۷/۵۳۴۷
SNDVI,۷							۰

۵-۲- انتخاب ویژگی با استفاده از الگوریتم ژنتیک

به منظور انتخاب ویژگی با استفاده از الگوریتم ژنتیک، تابع آمیزش scattered و تابع جهش، گوسین در نظر گرفته شد. تابع جهش گوسین مقداری که از توزیع گوسین با میانگین صفر، به صورت رندم انتخاب می‌شود به هر کدام از مختصات‌های والدین اضافه می‌کند. انحراف معیار این توزیع با استفاده از دو پارامتر scale و shrink تعیین می‌شود. پارامتر scale برابر با انحراف معیار توزیع در اولین نسل است و پارامتر shrink میزان تغییر (افزایش یا کاهش) انحراف معیار را در هر نسل تعیین می‌کند. در صورتی که مقدار shrink برابر با یک در نظر گرفته شود، مقدار انحراف معیار در هر نسل به طور خطی کاهش می‌یابد و در تکرار آخر مقدار آن به صفر می‌رسد (رابطه‌ی ۶)

$$\sigma_k = \sigma_{k-1}\left(1 - shrink\frac{k}{Generations}\right) \qquad (۶)$$

دو پارامتر scale و shrink برابر با ۱ و اندازه‌ی جمعیت برابر با ۳ و تعداد نسل‌ها برابر با ۱۰۰ در نظر گرفته شد. الگوریتم ژنتیک یک الگوریتم جستجوی تصادفی است و برای این که به نتیجه‌ی مطلوب دست یابد باید فضای جستجو به اندازه‌ی کافی بزرگ در نظر گرفته شود. به همین دلیل باید اندازه‌ی جمعیت یا تعداد نسل‌ها بالا در نظر گرفته شود که در این جا تعداد نسل‌ها برابر با ۱۰۰ در نظر گرفته شده و از افزایش تعداد جمعیت به دلیل افزایش سرعت الگوریتم اجتناب شد. نوع افراد[۱] در جمعیت از نوع رشته‌های بیتی[۲]تعریف شد. مقدار ۱ در این رشته‌ها به معنی انتخاب ویژگی مورد نظر و ۰ به معنی عدم انتخاب آن ویژگی است. طول هر کدام از افراد برابر با تعداد ویژگی‌های ورودی به فرایند انتخاب ویژگی در نظر گرفته شد. بنابراین طول رشته‌ی بیتی برای ویژگی‌های غیر حرارتی برابر با ۳۹ و در مورد ویژگی‌های حرارتی برابر با ۱۲ در نظر گرفته شد. تابع هدف در الگوریتم ژنتیک برابر با میزان عدم توافق نتیجه‌ی طبقه‌بندی به روش SVM با ماسک ابر مادیس تعریف شد. به منظور محاسبه‌ی تابع هدف ابتدا ۳۰ درصد از داده‌های آموزشی به صورت رندم به عنوان تست انتخاب شده و محاسبه‌ی مقادیر تابع هدف روی این داده‌های. تست صورت گرفت. شکل ۳ نمودار مقادیر تابع هدف (میزان fiteess) را در برابر تعداد نسل‌ها در طبقه‌بندی با ویژگی‌های انعکاسی و حرارتی نشان می‌دهد. نقاط آبی رنگ میانگین مقادیر تابع هدف در هر جمعیت و نقاط سیاه بهترین مقدار تابع هدف در نسل کنونی را نشان می‌دهد. با توجه به شکل مشاهده می‌شود در تکرارهای نهایی در مورد هر دو نمودار مقدار میانگین تابع هدف برابر با بهترین مقدار شده (انطباق نقاط آبی و مشکی) و مسئله همگرا می‌شود. در واقع در تکرارهای نهایی افراد موجود در جمعیت یکسان خواهند بود که با آمیزش این افراد تکراری دوباره همان افراد ایجاد خواهد شد و بهبودی در نتایج از این مرحله به بعد حاصل نخواهد شد. در نهایت در مورد ویژگی‌های حرارتی مقدار تابع هدف به مقدار ۱۱٪ و در مورد ویژگی‌های انعکاسی به مقدار ۱۲٪رسید.



شکل۳- انتخاب ویژگی با استفاده از الگوریتم ژنتیک؛ بالا: انتخاب ویژگی‌های حرارتی؛ پایین: ویژگی‌های غیر حرارتی

۱ individual
۲ bitstring

۵-۳- انتخاب ویژگی با استفاده از ماتریس کارلیشن

یکی از تابع‌های ارزیابی موجود در فرایند انتخاب ویژگی معیار وابستگی است. با استفاده از این معیار می‌توان ویژگی‌های اضافی که اطلاعات جدیدی در مورد کلاس‌ها به دست نمی‌دهند را حذف و افزونگی داده را کاهش داد. در این روش ابتدا بین ویژگی‌های ورودی، ماتریس کارلیشن محاسبه شده و سپس ویژگی‌هایی که دارای کارلیشن بیش از مقدار حدآستانه هستند از بین ویژگی‌های انتخابی حذف می‌شوند. مقدار حدآستانه در این روش برابر با ۰/۷۵ در نظر گرفته شد. در جداول ۴ و ۵ بخشی از ماتریس‌های کارلیشن برای ویژگی‌های غیرحرارتی و حرارتی نشان داده شده‌است. با توجه به این جداول ملاحظه می‌شود که ویژگی‌های واریانس ۱ و کنتراست ۱ دارای کارلیشن بالاتر از مقدار حدآستانه می‌باشند بنابراین یکی از این دو ویژگی از بین ویژگی‌های انتخابی حذف می‌شوند. همچنین ویژگی‌های میانگین باند ۳۱ و ۲۲ نیز جزو ویژگی‌های با کارلیشن بالا (۹۵٪) هستند و در نظر گرفتن هر دو ویژگی ذکر شده می‌تواند موجب کاهش دقت عملکرد طبقه‌بندی کننده شده و یکی از این دو ویژگی باید در نظر گرفته شود.

جدول۴- بخشی از ماتریس کارلیشن برای ویژگی‌های انعکاسی

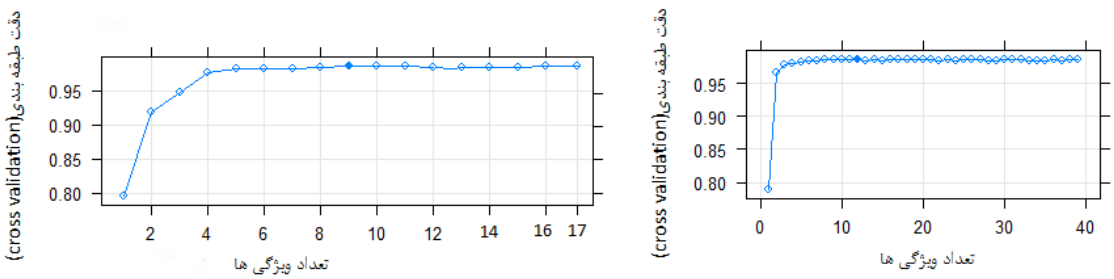
	NDVI	NDSI	B۱/B۲	mean۱	var۱	homo۱	contrast۱	mean۲	variance۲
NDVI	۱	۱۳/-۰	۹۲/-۰	۷۳۵/-۰	۳۲/-۰	۴۲۴/۰	۲۲/-۰	۶۲/-۰	۲۳/-۰
NDSI		۱	۰/۷	۰۴/-۰	۲۳۶/۰	۲۵/-۰	۲۳۶/۰	۲۴/-۰	۲۳/۰
B۱/B۲			۱	۲۳/-۰	۰۳/-۰	۱۴/۰	۰۳/-۰	۴۴/-۰	۰۳/۰
mean۱				۱	۳۴/۰	۶۴/-۰	۳۲/۰	۹۸/۰	۳۳/۰
var۱					۱	۵۲/-۰	۷۹/۰	۲۷/۰	۹۹/۰
homo۱						۱	۵۳/-۰	۵۴/-۰	۵۲/-۰
contrast۱							۱	۲۵۴/۰	۷۹/۰
mean۲								۱	۲۶/۰
variance۲									۱

جدول۵- بخشی از ماتریس کارلیشن برای ویژگی‌های حرارتی

	temp۳۲	bt۳۱-bt۳۲	bt۳۳-bt۳۱	bt۳۱-bt۲۰	bt۲۰-bt۳۲	corr۲۰	mean۲۲	second moment۲۲	mean۳۱
temp۳۲	۱	۰۶/-۰	۷/-۰	۰۹/-۰	۰۸۱/۰	۱۵/-۰	۸۳/۰	۰۷/-۰	۷۶/۰
bt۳۱-bt۳۲		۱	۱۱/۰	۱۹/-۰	۲۷/۰	۰۰۳/۰	۱۷/-۰	-۰۲	۲۶/-۰
bt۳۳-bt۳۱			۱	۲۲/-۰	۲۲/۰	۱۴/-۰	۹۱/-۰	۳۹/-۰	۹۴/-۰
bt۳۱-bt۲۰				۱	۹۹/-۰	۱۹/۰	۰۰۹/-۰	۲۶/۰	۳۰/۰
bt۲۰-bt۳۲					۱	۱۹/-۰	۰۰۵/-۰	۲۷/-۰	۳۲/-۰
corr۲۰						۱	۰۴۴/-۰	۴۱/۰	۰۲۹/۰
mean۲۲							۱	۲۵/۰	۹۵/۰
second moment۲۲								۱	۰/۳۳
mean۳۱									۱

۵-۴- انتخاب ویژگی با استفاده از حذف ویژگی به روش بازگشتی

در این روش، ابتدا طبقه‌بندی با استفاده از همه‌ی ویژگی‌ها انجام شده و برای هر ویژگی یک مقدار اهمیت (رتبه) محاسبه می‌شود. سپس طبقه‌بندی با استفاده از مهم‌ترین ویژگی انجام شده و دقت طبقه‌بندی محاسبه می‌شود این کار برای تعداد-های مختلف از مهم‌ترین ویژگی‌ها تکرار می‌شود. در نهایت آن تعداد از مهم‌ترین ویژگی‌ها که دارای بالاترین دقت طبقه‌بندی هستند، انتخاب می‌شوند. در این تحقیق به منظور پیاده سازی الگوریتم بازگشتی مدل طبقه‌بندی الگوریتم جنگل تصادفی در نظر گرفته شده و با استفاده از معیار اهمیت ویژگی، اهمیت هر ویژگی محاسبه می‌شود. به منظور پیاده سازی این روش نیز مانند روش ماتریس کارلیشن از بسته‌ی caret نرم‌افزار R استفاده شد. دقت مدل طبقه‌بندی با استفاده از روش cross 10-fold validation محاسبه شده و در انتها زیر مجموعه‌ای از ویژگی‌ها که بالاترین دقت را نتیجه می‌دهند، انتخاب می‌شوند. شکل ۴ دقت‌های به دست آمده را برحسب مقادیر مختلف تعداد ویژگی‌ها نشان می‌دهد. همان طور که در شکل ۴ دیده می‌شود بالاترین دقت در مورد ویژگی‌های انعکاسی با استفاده از ۱۲ ویژگی و برای ویژگی‌های حرارتی با استفاده از ۹ ویژگی به دست آمده است.



شکل۴- انتخاب ویژگی با استفاده از حذف ویژگی به روش بازگشتی؛ راست: ویژگی‌های غیر حرارتی؛ چپ: ویژگی‌های حرارتی.

۵-۵- طبقه‌بندی با استفاده از روش‌های یادگیری جمعی

همان طور که گفته شد در این تحقیق از ۴ نوع روش boosting شامل adaboost.M1، adaboostSVM، logitboost، totalboost و الگوریتم جنگل تصادفی به منظور طبقه‌بندی استفاده شد. طبقه‌بندی کننده‌ی SVM با کرنل گوسین در سه روش adaboost.M1، adaboostSVM و totalboost به عنوان طبقه‌بندی کننده‌ی پایه[۱] در نظر گرفته شد. از آن جایی که طبقه‌کننده‌ی SVM روشی پارامتریک است نیاز است قبل از انجام طبقه‌بندی پارامترهای مورد نظر تنظیم شوند. پارامترهایی که باید در طبقه‌بندی با SVM-RBF تنظیم شوند شامل دو پارامتر هزینه (C) و گاما (γ) می‌باشند. به منظور تنظیم این پارامترها از روش جستجوی گرید[۲] استفاده شد. در این روش با در نظر گرفتن محدوده‌ای از مقادیر C و γ مقادیری که بالاترین دقت cross validation را نتیجه می‌دهند به عنوان پارامترهای نهایی در نظر گرفته می‌شوند. جستجوی مقدار بهینه برای پارامتر C در محدوده‌ی 2^{-5} تا 2^{15} و برای پارامتر γ از 2^{-15} تا 2^{3} انجام و تعداد fold ها برابر با ۵ در نظر گرفته شد. در هر تکرار از روش‌های adaboost.M1 و totalboost مقادیر پارامترهای بهینه با روش جستجوی گرید انتخاب می‌شوند. به منظور افزایش سرعت در اجرای الگوریتم‌های boosting در روش‌هایی که تعداد دفعات تکرار از پیش قابل تعیین است (روش‌های adaboost.M1، logitboost و totalboost) تعداد دفعات تکرار برابر با ۳ در نظر گرفته شد.

۵-۵-۱- طبقه‌بندی با استفاده از روش adaboost.M1

در روش adaboost.M1 در هر تکرار یک زیر مجموعه از داده‌های آموزشی به طور تصادفی انتخاب شده و توزیع وزن داده‌های آموزشی به گونه‌ای بروزرسانی می‌شود که وزن بیش‌تری به داده‌های آموزشی که اشتباه طبقه‌بندی شده‌اند اختصاص یابد. شکل‌های ۱۱ الف و ب نقشه‌های طبقه‌بندی به دست آمده با استفاده از ادغام با این روش را نشان می‌دهند، که به ترتیب از روش‌های معیار S و GA به منظور انتخاب ویژگی استفاده شده است.

۵-۵-۲- طبقه‌بندی با استفاده از روش adaboostSVM

همان طور که توضیح داده شد در روش adaboostSVM مقدار پارامتر هزینه (C) ثابت نگه داشته شده و مقدار σ در صورتی که خطای طبقه‌بندی بزرگ‌تر از ۰/۵ شود، به اندازه‌ی σ_{scale} کاهش می‌یابد و الگوریتم زمانی متوقف می‌شود که مقدار σ برابر با σ_{min} شود. علت توقف الگوریتم در σ_{min} این است که انتخاب مقادیر بسیار کوچک برای σ باعث انطباق بیش از حد مدل طبقه‌بندی به داده‌های آموزشی می‌شود. در این تحقیق مقدار پارامتر هزینه (C) در روش adaboostSVM برابر با 2^{15} در نظر گرفته شد. از آن جایی که انتخاب مقادیر کوچک برای پارامتر هزینه می‌تواند باعث کاهش دقت طبقه‌بندی شود باید از انتخاب مقادیر کوچک اجتناب شود. برای طبقه‌بندی با استفاده از این روش نیاز به تعیین پارامترهای $\sigma_{initial}$، σ_{min} و σ_{scale} می‌باشد. هر چند جواب نهایی وابسته به مقادیر $\sigma_{initial}$ و σ_{scale} نمی‌باشد، اما در نظر گرفتن σ_{scale} مناسب می‌تواند تعداد دفعات تکرار الگوریتم را کاهش و موجب سریع‌تر همگرا شدن مقدار $\sigma_{initial}$ به σ_{min} شود. در این تحقیق مقدار $\sigma_{initial}$ (پارامتر سیگمای کرنل گوسین در تکرار اول) برابر با شعاع پراکندگی داده‌های آموزشی در فضای ویژگی‌ها و مقدار σ_{min} (پارامتر سیگمای کرنل گوسین در تکرار آخر) برابر با میانگین مقادیر کم‌ترین فاصله بین هر دو داده‌ی آموزشی در نظر گرفته شدند[۲۰]. شکل‌های ۱۱ ج و د نقشه‌های طبقه‌بندی به دست آمده با استفاده از ادغام با این روش را نشان می‌دهند، که در این شکل‌ها، به ترتیب از روش‌های معیار S و GA به منظور انتخاب ویژگی استفاده شده است.

[۱] Base learner
[۲] Grid search

۵-۵-۳- طبقه‌بندی با استفاده از روش logitboost

همان طور که گفته شد در روش لاجیت بوست در هر تکرار به مقادیر پاسخ (zj) هر کلاس یک مدل برازش داده می‌شود. مدل‌های برازش داده شده به هر کلاس، در هر تکرار، با یکدیگر جمع می‌شوند و از مقدار به دست آمده به منظور به روز رسانی احتمال تعلق هر پیکسل به هر کلاس استفاده می‌شود و دوباره در تکرار بعد با استفاده از مقادیر احتمال جدید مقدار z_j محاسبه شده و این فرایند تکرار می‌شود. خروجی این روش به صورت مقادیر احتمال است. در این تحقیق به منظور برازش مدل به هر کلاس در هر تکرار، از یک مدل خطی از ویژگی‌های ورودی با ترم ثابت، استفاده شد. تعداد ترم‌های مدل (تعداد ویژگی‌ها) با استفاده از الگوریتم ژنتیک برای ویژگی‌های انعکاسی و حرارتی بهینه شدند. تابع هدف مشابه سایر روش‌ها می‌باشد با این تفاوت که به جای طبقه‌بندی کننده‌ی SVM از روش logit boost استفاده شد.

۵-۵-۴- طبقه‌بندی با استفاده از روش totalboost

همان طور که توضیح داده شد، در روش totalboost در هر تکرار وزن داده‌های آموزشی به گونه‌ای به روز رسانی می‌شوند که مقدار آنتروپی نسبی نسبت به توزیع اولیه‌ی داده‌های آموزشی مینیمم شود. این مسئله‌ی بهینه سازی با استفاده از روش برنامه‌سازی درجه دو مقید حل می‌شود و دارای دو قید است: ۱- مجموعه مقادیر اضافه شده به وزن هر کدام از داده‌های آموزشی برابر با صفر شود. ۲- مقادیر edge ها (میانگین وزن دار مقادیر حاشیه‌ها) در تکرار فعلی و تکرارهای قبل کوچک‌تر مساوی با $\min(\gamma_q)\text{-v } q=1,\ldots,t$ (اختلاف پارامتر دقت (v) از مینیمم مقادیر edge ها در t تکرار قبل) شود. به منظور پیاده‌سازی پارامتر دقت (v) برابر با ۰/۰۲ نظر گرفته شد. شکل ۱۲ ب نقشه‌ی طبقه‌بندی حاصل از ادغام با استفاده از این روش را نشان می‌دهد.

جدول۶- ویژگی‌های انعکاسی و حرارتی انتخاب شده با استفاده از روشهای معیار S، الگوریتم ژنتیک (GA)، ماتریس کارلیشن و روش بازگشتی

روش انتخاب ویژگی	ویژگی‌های انعکاسی		ویژگی‌های حرارتی	
	ویژگی‌های بافت	*ویژگی‌های طیفی*	*ویژگی‌های بافت*	*ویژگی‌های طیفی*
معیار S	میانگین باند ۱ عدم شباهت باند ۲ آنتروپی باند۲ گشتاور دوم باند ۲	انعکاس باند ۱ انعکاس باند ۸ میانگین انعکاس باند ۸ انعکاس باند ۲۶ NDSI، نسبت انعکاس باند ۱ به باند ۲	میانگین باند ۳۱	دمای باند ۳۲و۳۵ $BT_{۳۳}$-$BT_{۳۱}$ $BT_{۳۱}$-$BT_{۲۰}$ $BT_{۲۰}$-$BT_{۳۲}$ $BT_{۲۲}$-$BT_{۳۲}$
GA (تابع هدف: نزدیکی طبقه‌بندی SVM به MOD35)	واریانس باند ۱ همگنی باند۱ آنتروپی باند۱ میانگین باند۲.	نسبت انعکاس باند ۱ به باند ۲، NDSI	عدم شباهت باند ۳۱، میانگین باند ۳۲، وابستگی باند۲۰، میانگین باند ۳۲	دمای باند ۳۵ $BT_{۳۳}$-$BT_{۳۱}$ $BT_{۳۱}$-$BT_{۲۰}$ $BT_{۲۰}$-$BT_{۳۲}$
GA (تابع هدف: نزدیکی طبقه‌بندی logitboost به MOD35)	واریانس ۱، میانگین ۲، کنتراست ۸، میانگین ۲۶	NDSI	کارلیشن ۲۰، میانگین ۲۲، گشتاور دوم ۲۲، میانگین ۳۱، میانگین ۳۲.	$BT_{۲۲}$-$BT_{۳}$ $BT_{۳۴}$-$BT_{۳۵}$
ماتریس کارلیشن	وابستگی ۱، کنتراست۲، عدم شباهت۲، گشتاور مرتبه دوم ۲، وابستگی ۲، واریانس ۸، وابستگی ۸، میانگین ۲۶، واریانس ۲۶ و وابستگی ۲۶	انعکاس ۲۶، نسبت انعکاس باند ۱ به ۲،NDVIو NDSI	وابستگی ۲۰، میانگین ۲۲، گشتاور مرتبه دوم ۲۲، عدم شباهت ۳۱	دمای باند ۲۰، ۲۲، ۳۱ و ۳۵ $BT_{۲۰}$-$BT_{۳۲}$ $BT_{۳۱}$-$BT_{۲۰}$ $BT_{۳۱}$- $BT_{۳۲}$
روش بازگشتی	، میانگین ۸، میانگین ۱، میانگین ۲، میانگین ۲۶، آنتروپی ۱	انعکاس باند۸، انعکاس ۲، انعکاس ۱، انعکاس ۲۶، NDSI، NDVI، نسبت انعکاس باند ۱ به ۲	میانگین ۳۲، میانگین ۲۲، میانگین ۳۱، عدم شباهت ۳۱	دمای باند ۳۲ $BT_{۳۳}$-$BT_{۳۱}$ $BT_{۲۰}$-$BT_{۳۲}$ $BT_{۳۱}$-$BT_{۲۰}$ $BT_{۳۱}$- $BT_{۳۲}$

۵-۵-۵- طبقه‌بندی با استفاده از الگوریتم RF

دو پارامتر باید پیش از طبقه‌بندی با الگوریتم جنگل تصادفی تعیین شوند. تعداد درخت‌ها و تعداد ویژگی‌های انتخاب شده در هر گره. در اکثر مطالعات پیشین گزارش شده است که خطا قبل از ۵۰۰ درخت به پایداری می‌رسد. همچنین در این تحقیقات دو مقدار برای تعداد ویژگی‌های انتخاب شده در هر گره در نظر گرفته شده است: ۱- کل تعداد ویژگی‌های ورودی ۲- مجذور تعداد کل ویژگی‌ها. مورد اول باعث افزایش زمان اجرای الگوریتم می‌شود، بنابراین تعداد ویژگی‌های انتخاب شده در هر گره در این مقاله برابر با جذر تعداد کل ویژگی‌ها (مقدار پیش-فرض در R) در نظر گرفته شد[۳۱].

شکل‌های ۱۲ ج و د نقشه‌های طبقه‌بندی حاصل از ادغام با استفاده از روش RF در صورتی که به ترتیب، از روش معیار S و GA به منظور انتخاب ویژگی استفاده شده باشد را نشان می‌دهند. همچنین شکل ۱۳ الف و ب نقشه‌های طبقه‌بندی مربوط به این روش می‌باشند، که از به ترتیب از روش‌های RFE و ماتریس کارلیشن به منظور انتخاب ویژگی استفاده شده باشد.

۵-۶- ارزیابی نتایج

به منظور مقایسه‌ی روش‌های مختلف طبقه‌بندی از نظر دقت کاپای هر کلاس که از ماتریس ابهام[۱] به دست می‌آیند، استفاده شده‌است. همچنین میزان توافق نتایج طبقه‌بندی در قسمت‌های ابری با نقشه‌ی مرجع به دست آمده از ماسک ابر مادیس (MOD35) محاسبه شد.

۵-۶-۱- ارزیابی عملکرد روش‌های یادگیری جمعی در شناسایی پیکسل‌های ابر، برف/یخ و سیروس با توجه به روش انتخاب ویژگی

در این قسمت عملکرد روش‌های یادگیری جمعی با توجه شاخص‌های مربوط به هر یک از کلاس‌های ابر، برف/یخ و سیروس (دقت‌های تولیدکننده و کاربری و ضریب کاپای هر کلاس) در حالتی که از روش‌های معیار S، الگوریتم ژنتیک (GA)، ماتریس کارلیشن و حذف ویژگی به روش بازگشتی (RFE) برای انتخاب ویژگی استفاده شده باشد، با یکدیگر مقایسه شده‌اند. به منظور بررسی عملکرد روش‌های مختلف انتخاب ویژگی و یادگیری جمعی در شناسایی پیکسل‌های ابری، برف/یخ و سیروس نیاز به بررسی دقت‌های تولید کننده، کاربری و کاپای هر کلاس می‌باشد. همان طور که در جدول ۸ مشاهده می‌شود، به طور کلی همه‌ی روش‌های یادگیری جمعی به دقت‌های تولید کننده‌ی بالایی در شناسایی کلاس ابر دست یافتند و به استثنای روش totalboost-GA برای همه‌ی روش‌ها نیز دقت تولید کننده‌ی بالایی در شناسایی ابرهای سیروس به دست آمده است. دقت‌های کاربری ابر به دست آمده در حالتی که از روش ماتریس کارلیشن و RFE در الگوریتم جنگل تصادفی به منظور انتخاب ویژگی استفاده شده، به ترتیب مقادیر ۱۰۰٪ و ۹۹٪ به دست آمد که نسبت به مقادیر مشابه برای روش‌های معیار S و الگوریتم ژنتیک (GA)، مقادیر بالاتری است. همچنین ملاحظه می‌شود که اکثر روش‌های boosting صرف نظر از این که از چه روشی برای انتخاب ویژگی استفاده شود، دقت های تولید کننده‌ی برف/یخ بالاتری نسبت به الگوریتم جنگل تصادفی نتیجه دادند.

از مزایای روش‌های boosting ایجاد توازن بین تعداد داده‌های آموزشی کلاس‌ها در صورتی است که تعداد داده‌های آموزشی یک کلاس اختلاف زیادی با تعداد داده‌های آموزشی کلاس‌های دیگر داشته باشد، این امر با اختصاص وزن بیش‌تر به داده‌های آموزشی مربوط به کلاس با داده‌های آموزشی کم‌تر، در صورت طبقه‌بندی نادرست داده‌های آموزشی آن کلاس میسر می‌شود. با توجه به جدول ۷ مشاهده می‌شود که تعداد داده‌های آموزشی پیکسل‌های برف/یخ نسبت به سایر کلاس‌ها کم‌تر است، در نتیجه در صورت طبقه‌بندی اشتباه پیکسل‌های برف/یخ در روش‌های boosting وزن بیش‌تری به داده‌های آموزشی مربوطه که اشتباه طبقه‌بندی شده‌اند اختصاص می‌یابد و در تکرارهای بعدی مدل طبقه‌بندی به گونه‌ای تنظیم می‌شود که داده‌ی آموزشی مورد نظر، در کلاس صحیح قرار گیرد. ولی در الگوریتم RF چنین قابلیتی وجود ندارد، در نتیجه دقت تولیدکننده‌ی کم‌تری برای پیکسل‌های برف/یخ با استفاده از روش RF به دست آمده است.

در روش‌های boosting دقت‌های کاربری ابر سیروس نسبتا بالاتری نسبت به اکثر حالت‌های الگوریتم جنگل تصادفی به دست آمده است. همچنین حذف ویژگی‌های

[۱] Confusion matrix

دارای وابستگی بیش از ۰/۷۵ در RF باعث بهبود دقت کاربری ابرهای سیروس شده‌است. با مقایسه‌ی روش‌هایی که از الگوریتم ژنتیک برای انتخاب ویژگی استفاده کرده‌اند، ملاحظه می‌شود که برخی از روش‌ها مانند RF توانسته‌اند به دقت کاربری بالایی در شناسایی پیکسل‌های برف/یخ دست یابند و روش logit boost-GA دقت بسیار پایین‌تری را در شناسایی پیکسل‌های برف/یخ نسبت به این روش‌ها نشان داده و نتوانسته است هیچ پیکسل برف/یخ را در تصویر شناسایی کند. علت این تفاوت به محدودیت‌های موجود در روش logit boost مربوط می‌شود، طبقه‌بندی کننده‌های پایه در هر تکرار، در سایر روش‌های SVM boosting می‌باشد، ولی در الگوریتم logit boost مدل طبقه‌بندی با استفاده از برازش یک مدل خطی در هر تکرار به دست می‌آید که نسبت به طبقه‌بندی کننده‌ی SVM دارای پیچیدگی کم‌تری است. از بین روش‌های boosting که از طبقه‌بندی کننده‌ی پایه‌ی SVM استفاده می‌کنند، روش adaboost.M1-GA میانگین دقت‌های تولیدکننده و کاربری بالاتری را در شناسایی پیکسل‌های برف/یخ نشان می‌دهد و روش‌های adaboostSVM-GA و totalboost-GA مقادیر مشابهی را نشان می‌دهند.در میان حالت‌های مختلف روش RF، روش RF-S پایین‌ترین میانگین دقت تولیدکننده و کاربری را در شناسایی پیکسل‌های برف/یخ، نشان می‌دهد.

با مقایسه‌ی مقدار ضریب کاپا برای کلاس ابر ملاحظه می‌شود که در روش‌های boosting مقدار ضریب کاپا بالاتر از ۰/۹۳ به دست آمده است که نشان دهنده‌ی قدرت بالای روش‌های boosting در شناسایی پیکسل‌های ابری است (شکل ۸). در روش RF بسته به نوع روش انتخاب ویژگی، مقدار ضریب کاپا برای کلاس ابر بین ۰/۶۵ تا ۱ متغیر بوده است. انتخاب ویژگی با استفاده از روش‌های ماتریس کارلیشن و بازگشتی، منجر به افزایش ضریب کاپا به حداکثر مقدار خود شده است. روش‌های GA و معیار S در صورت ثابت بودن روش ادغام، مقادیر کاپای تقریبا مشابهی برای کلاس ابر نتیجه دادند.

با توجه به شکل ۸ مشاهده می‌شود که بالاترین ضریب کاپای کلاس برف/یخ در روش‌های boosting مربوط به روش adaboostSVM-S و در روش RF مربوط به روش ماتریس کارلیشن می‌باشد. پایین‌ترین ضریب کاپای کلاس برف/یخ در روش‌های boosting که از طبقه‌بندی کننده‌ی پایه‌ی SVM استفاده می‌کنند، همان طور که در مورد دقت کاربری مشاهده شد، مربوط به روش totalboost-GA و در الگوریتم RF مربوط به حالتی است که از روش معیار S برای انتخاب ویژگی استفاده شده است. روش logit boost همان طور که در مورد دقت کاربری ذکر شد، نتوانست پیکسل برف/یخ در تصویر شناسایی کند و ضریب کاپا برای آن تعریف نمی‌شود. سایر حالت‌های روش‌های boosting و RF ضریب کاپای کلاس برف/یخ را به ترتیب بین ۰/۷-۰/۵ و ۰/۵-۰/۴ نتیجه می‌دهند. با تغییر روش انتخاب ویژگی از معیار S به روش الگوریتم ژنتیک در روش adaboostSVM ملاحظه می‌شود که مقدار ضریب کاپای کلاس برف/یخ ۱۸٪ کاهش یافت که دو برابر مقدار مشابه برای روش adaboost.M1 است. این موضوع نشان دهنده‌ی حساسیت بالای روش adaboostSVM نسبت به روش انتخاب ویژگی است. روش adaboostSVM همان طور که در مورد کلاس سیروس هم دیده می‌شود (شکل ۸)، حساسیت بیش‌تری نسبت به روش adaboost.M1 به ویژگی‌های ورودی دارد. دو پارامتر σ_{ini} و σ_{min} در روش adaboostSVM با استفاده از شعاع پراکندگی داده‌های آموزشی در فضای ویژگی تعیین می‌شوند و وابستگی روش adaboostSVM به مقدار σ_{min} باعث حساسیت بیش‌تر این روش به ویژگی‌های ورودی، نسبت به adaboost.M1 می‌شود.

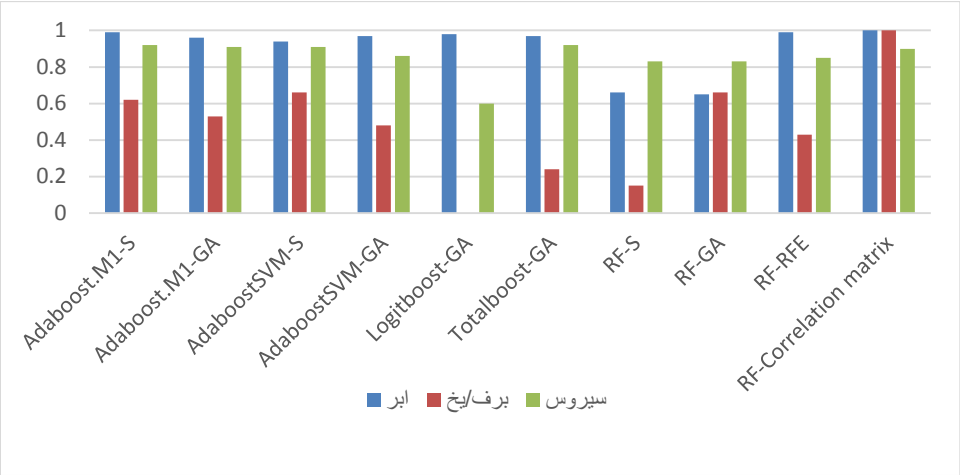
ضریب کاپای به دست آمده برای کلاس سیروس برای اکثر روش‌های یادگیری جمعی آورده شده در جدول ۸، مقداری بالاتر از ۰/۸ را نشان می‌دهد که به معنی قابلیت بالای این روش‌ها در شناسایی پیکسل‌های سیروس است (شکل ۸). از بین روش‌های انتخاب ویژگی استفاده شده در RF، روش ماتریس کارلیشن، ضریب کاپای بزرگ‌تری برای کلاس سیروس نسبت به سایر روش‌های انتخاب ویژگی نتیجه داده است. نکته‌ی قابل توجه در مورد کلاس سیروس این است که کم‌ترین دقت‌های کاربری مربوط به انتخاب ویژگی با استفاده از روش GA می‌باشد، در صورتی که به ویژگی‌های انتخابی در الگوریتم ژنتیک توجه شود، مشاهده می‌شود که هیچ یک از ویژگی‌های به کار رفته در ماسک ابر مادیس که برای شناسایی پیکسل‌های سیروس مورد استفاده قرار می‌گیرند، انتخاب نشده‌اند. در نتیجه دقت کاربری به دست آمده برای روش adaboost.M1، adaboostSVM و RF نسبت به حالتی که از روش معیار S استفاده شده، به ترتیب ۰/۱٪، ۰/۴٪ و ۰/۱٪ کاهش داشته است.

جدول۷- تعداد نمونه‌های آموزشی و داده‌های تست

شماره کلاس	نام کلاس	تعداد داده‌های آموزشی(پیکسل)	تعداد داده‌های تست(پیکسل)
۱	ابر	۹۸۸	۱۰۹
۳	برف/یخ	۱۱۴	۱۲
۶	ابر سیروس	۵۵۱	۶۱

جدول۸- مقادیر دقت تولید کننده و کاربری برای انواع مختلف روش‌های boosting و الگوریتم RF با روش‌های انتخاب ویژگی مختلف

روش	برچسب کلاس	کلاس	دقت تولید کننده	دقت کاربری
Adaboost.M1-S	۱	ابر	٪۱۰۰	٪۹۹
	۳	برف/یخ	٪۵۸	٪۶۴
	۶	ابر سیروس	٪۹۳	٪۹۳
Adaboost.M1-GA	۱	ابر	٪۹۹	٪۹۷
	۳	برف/یخ	٪۵۰	٪۵۵
	۶	ابر سیروس	٪۹۷	٪۹۲
adaboostSVM-S	۱	ابر	٪۱۰۰	٪۹۶
	۳	برف/یخ	٪۵۰	٪۶۷
	۶	ابر سیروس	٪۹۵	٪۹۲
AdaboostSVM-GA	۱	ابر	٪۹۹	٪۹۸
	۳	برف/یخ	٪۴۲	٪۵۰
	۶	ابر سیروس	٪۱۰۰	٪۸۸
Logitboost-GA	۱	ابر	٪۵۸	٪۹۸
	۳	برف/یخ	۰	-
	۶	ابر سیروس	٪۸۸	٪۹۷
Totalboost-GA	۱	ابر	٪۹۹	٪۹۲
	۳	برف/یخ	٪۶۷	٪۲۷
	۶	ابر سیروس	٪۶۷	٪۹۳
RF-GA	۱	ابر	٪۹۹	٪۷۵
	۳	برف/یخ	٪۳۳	٪۶۷
	۶	ابر سیروس	٪۹۷	٪۸۵
RF-RFE	۱	ابر	٪۱۰۰	٪۹۹
	۲	برف/یخ	٪۳۳	٪۴۴
	۶	ابر سیروس	٪۹۲	٪۸۷
RF-S	۱	ابر	٪۹۹	٪۷۶
	۳	برف/یخ	٪۴۲	٪۱۸
	۶	ابر سیروس	٪۱۰۰	٪۸۶
RF-correlation matrix	۱	ابر	٪۱۰۰	٪۱۰۰
	۳	برف/یخ	٪۲۹	٪۱۰۰
	۶	ابر سیروس	٪۱۰۰	٪۹۱



شکل۸- مقادیر ضریب کاپای کلاس‌های ابر، برف/یخ و سیروس با توجه به روش انتخاب ویژگی در انواع مختلف boosting و RF

۵-۶-۲- مقایسه‌ی روش‌های مختلف انتخاب ویژگی از نظر میزان نزدیکی با نقشه‌ی مرجع

از ماسک ابر مادیس در کاربردهای زیادی استفاده شده است. برخی از محققین از MOD35 به منظور راستی آزمایی روش‌های پیشنهادی خود استفاده کرده‌اند[۳]. بنابراین در این تحقیق از ماسک ابر مادیس به عنوان مرجع استفاده شد. ماسک ابر مادیس از ۱۹ باند از تصاویر مادیس به منظور بررسی ابری بودن پیکسل‌ها استفاده می‌کند. با استفاده از این باندها مقادیر انعکاس، نسبت‌های انعکاسی، دما و اختلاف دما محاسبه شده و بر روی آن‌ها حدآستانه تعریف می‌شود. نتایج بررسی‌ها پیکسل‌ها را به ۴ سطح اطمینان، بدون ابر(clear)، احتمالا بدون ابر(probably clear)، احتمالا ابری (probably cloud) و ابری(cloudy)، تقسیم می‌کند. این محصول از ۴۸ بیت تشکیل شده است که در آن علاوه بر اطلاعات در مورد موانع سطح، سایر اطلاعات کمکی که برای به دست آوردن پیکسل‌های ابری مورد نیاز است، فراهم شده است. همچنین شامل تست‌های طیفی است که برای به دست آوردن سطوح اطمینان از آن‌ها استفاده شده است. شکل ۹ فرایند استخراج یک نقشه‌ی مرجع که قابلیت مقایسه با نقشه‌های طبقه‌بندی به دست آمده از روش‌های یادگیری جمعی را داشته باشد، نشان می‌دهد. شکل ۱۰ نقشه‌ی مرجع به دست آمده از ماسک ابر مادیس را نشان می‌دهد.

در جدول ۹ درصد توافق روش‌های ذکر شده در جدول ۸ با ماسک ابر مادیس آورده شده است. با توجه به جدول ملاحظه می‌شود که روش‌های RF صحت بالاتری در شناسایی ابر (ابر سیروس و سایر ابرها) در مقایسه با روش‌های boosting نتیجه می‌دهند. روش RF-RFE همان طور که بر روی داده‌های تست دقت کاربری و ضریب کاپای بالایی نتیجه داد، به درصد توافق ۷۶٪ با نقشه‌ی مرجع در قسمت‌های ابری رسید. پایین‌ترین درصد توافق مربوط به روش logitboost-GA می‌باشد. با تغییر روش انتخاب ویژگی در adaboost.M1 و adaboostSVM از معیار S به روش GA صحت شناسایی نواحی ابری ۲ درصد بهبود یافته است.

۶- نتیجه گیری و پیشنهادات

در این تحقیق هدف شناسایی روش انتخاب ویژگی مناسب در ادغام طبقه‌بندی کننده‌های انعکاسی و حرارتی به منظور شناسایی پیکسل‌های ابر، برف/یخ و سیروس بود. اکثر روش‌های boosting صرف نظر از روش انتخاب ویژگی، دقت‌های تولید کننده و کاربری بالایی در شناسایی پیکسل‌های ابری و سیروس به دست آوردند. حذف ویژگی‌های افزونه در الگوریتم RF باعث افزایش دقت کاربری پیکسل‌های ابری شد. همچنین مقادیر ضریب کاپای به دست آمده برای هر ۳ کلاس، با استفاده از حذف ویژگی‌های افزونه نسبت به سایر روش‌های انتخاب ویژگی در RF بهبود یافت. بنابراین روش ماتریس کارلیشن عملکرد بهتری نسبت به سایر روش‌ها داشته است.

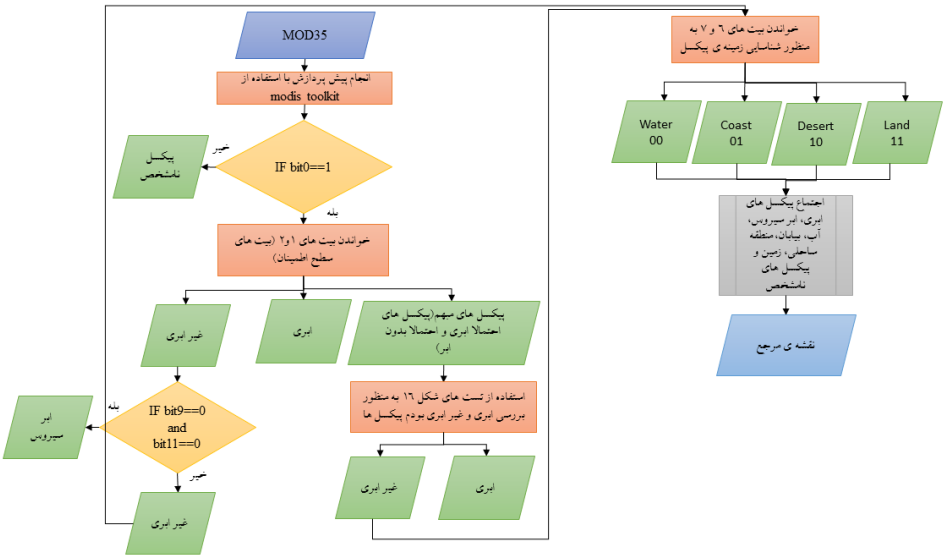
همان طور که گفته شد، یکی از مزایای روش‌های boosting ایجاد توازن بین کلاس‌ها با تعداد داده‌های آموزشی متفاوت است. در صورتی که داده‌های آموزشی یک کلاس نسبت به کلاس‌های دیگر کم‌تر باشد، این اختلاف با در نظر گرفتن وزن بیش‌تر برای داده‌های آموزشی مربوط به کلاس با تعداد داده‌های آموزشی کم‌تر جبران می‌شود. در حالت خاص که طبقه‌بندی کننده‌ی پایه SVM باشد، وزن داده‌ی آموزشی در پارامتر هزینه (C) ضرب می‌شود که با افزایش وزن، تاثیر آن داده‌ی آموزشی در ساخت مدل

طبقه‌بندی افزایش می‌یابد. همان طور که در مورد پیکسل‌های برف/یخ مشاهده شد، این قابلیت باعث افزایش میانگین دقت تولیدکننده‌ی پیکسل‌های برف/یخ در روش‌های boosting نسبت به روش RF شد.

همان طور که دیده شد، روش adaboostSVM به دلیل وابستگی پارامتر σ_{min} به ویژگی‌های ورودی حساسیت بیش‌تری نسبت به روش adaboost.M1 در شناسایی پیکسل‌های برف/یخ و سیروس نسبت به ویژگی‌های ورودی نشان داد.

الگوریتم RF نسبت به روش‌های boosting صرف نظر از روش انتخاب ویژگی، توافق بیش‌تری با ماسک ابر مادیس نشان داد. هر چند ویژگی‌های حرارتی و انعکاسی انتخابی در الگوریتم ژنتیک شامل ویژگی‌های مفید برای شناسایی ابر-های سیروس نبودند، به دلیل نزدیکی بالای نقشه‌ی طبقه-بندی به دست آمده در این روش، در سایر قسمت‌های ابری با نقشه‌ی مرجع، این موضوع بر روی میزان توافق تاثیر چندانی نداشته و روش adaboostSVM-GA بالاترین میزان توافق را از بین روش‌های boosting نتیجه داد. از بین روش-های انتخاب ویژگی به کار برده شده در RF روش حذف ویژگی به روش بازگشتی (RFE) توانست بالاترین میزان توافق را در بین روش‌های RF و boosting به خود اختصاص دهد. از بین روش‌های انتخاب ویژگی، روش‌های معیار S، GA، ماتریس کارلیشن وRFE به ترتیب ۱۰، ۶، ۵ و ۱۱ ویژگی از ویژگی‌های به کار گرفته شده در ماسک ابر مادیس را انتخاب کردند. بنابراین روش RFE که بیش‌ترین تعداد ویژگی‌های مشترک با ماسک ابر مادیس را دارد، بیش‌ترین نزدیکی با ماسک ابر مادیس را به دست آورد.

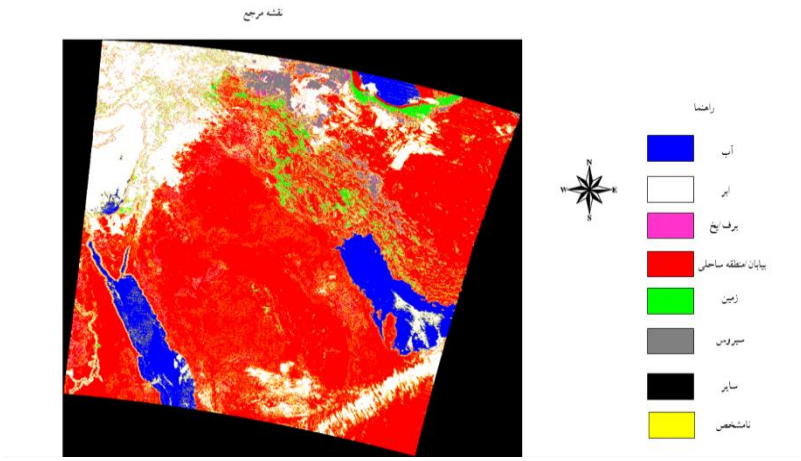
در این تحقیق ۱۰٪ از داده‌های آموزشی به صورت رندم به عنوان داده‌ی تست انتخاب شدند، افزایش داده-های تست می‌تواند روی نتایج نهایی تاثیر گذار باشد. همچنین می‌توان علاوه بر بررسی عملکرد روش‌های مختلف انتخاب ویژگی، اثر تغییر نوع ویژگی ورودی (بافتی یا طیفی) روی نتایج نهایی مورد بررسی قرار گیرد.



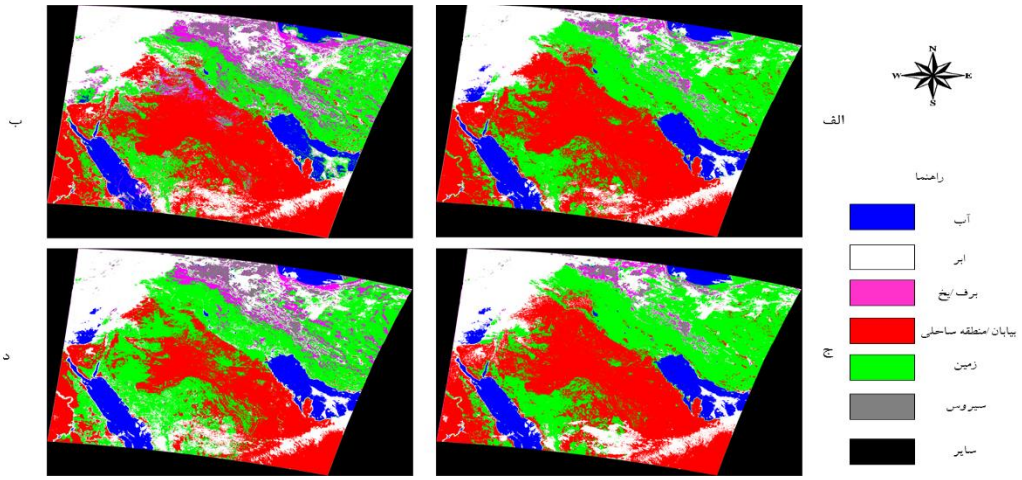
شکل۹- مراحل ایجاد نقشه‌ی مرجع با استفاده از ماسک ابر مادیس (MOD35)

جدول۹- درصد توافق روش‌های یادگیری جمعی با توجه به روش انتخاب ویژگی با نقشه‌ی مرجع به دست آمده از MOD35

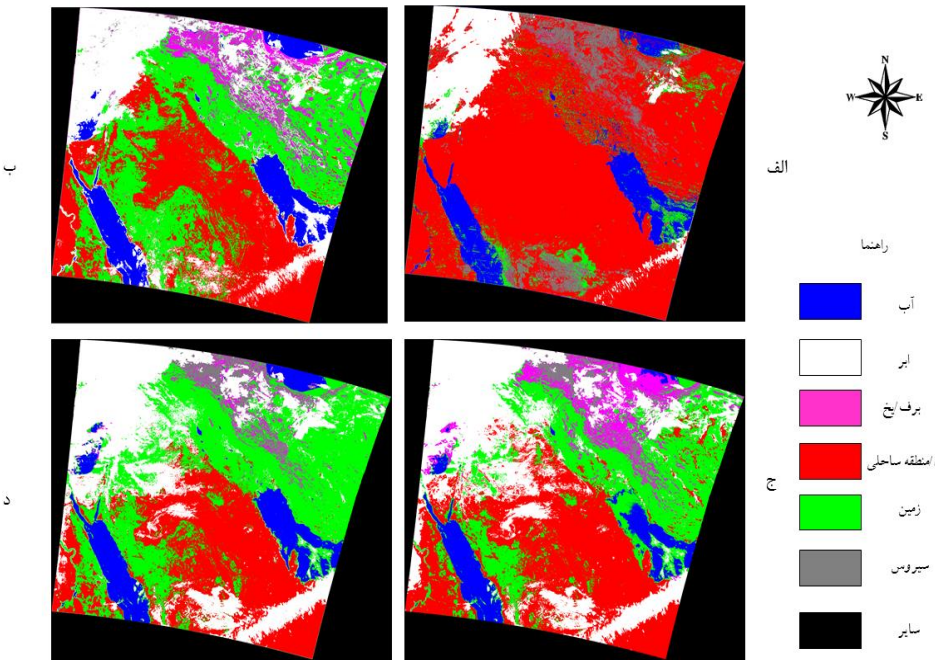
روش طبقه‌بندی و انتخاب ویژگی	درصد توافق در قسمت‌های ابری با نقشه‌ی مرجع
Adaboost.M1-S	٪۶۷
Adaboost.M1-GA	٪۶۹
AdaboostSVM-S	٪۶۹
Adaboost.SVM-GA	٪۷۱
Totalboost-GA	٪۶۳
Logitboost-GA	٪۴۲
RF-GA	٪۷۵
RF-RFE	٪۷۶
RF-S	۷۵٪
RF-correlation matrix	٪۷۴



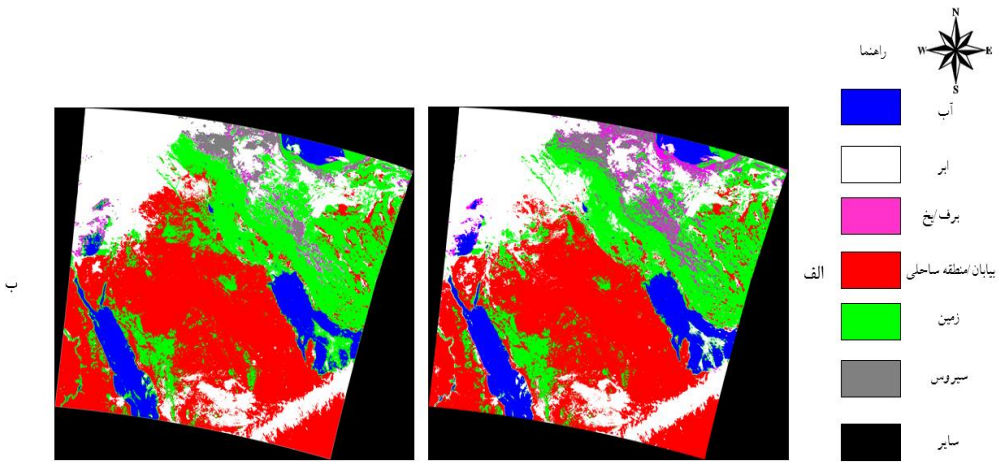
شکل۱۰- نقشه‌ی مرجع به دست آمده با استفاده از ماسک ابر مادیس (MOD35)



شکل۱۱- طبقه‌بندی با استفاده از روش‌های adaboost.M1 و adaboostSVM؛ الف: adaboost.M1-S؛ ب: adaboost.M1-GA؛ ج: adaboostSVM-S؛ د: adaboostSVM-GA



شکل۱۲- طبقه‌بندی با استفاده از روش‌های logit boost، totalboost و الگوریتم جنگل تصادفی؛ الف: logit boost-GA ؛ ب: totalboost-GA؛ ج: RF-S؛ د: RF-GA



شکل۱۳– طبقه‌بندی با استفاده از الگوریتم جنگل تصادفی؛ الف: RF-RFE ؛ ب: RF-correlation matrix.

مراجع

[1] Dash, M. and Liu, H., 1997. Feature selection for classification. Intelligent data analysis, 1(1-4), pp.131-156.

[2] Han B, Kang L, Song H. A fast cloud detection approach by integration of image segmentation and support vector machine. InInternational Symposium on Neural Networks 2006 May 28 (pp. 1210-1215). Springer Berlin Heidelberg.

[3] Guo F, Shen X, Zou L, Ren Y, Qin Y, Wang X, Wu J. Cloud Detection Method Based on Spectral Area Ratios in MODIS Data. Canadian Journal of Remote Sensing. 2015 Nov 2;41(6):561-76.

[4] Tang H, Yu K, Hagolle O, Jiang K, Geng X, Zhao Y. A cloud detection method based on a time series of MODIS surface reflectance images. International Journal of Digital Earth. 2013 Dec 9;6(sup1):157-71.

[5] Leinenkugel P, Kuenzer C, Dech S. Comparison and enhancement of MODIS cloud mask products for Southeast Asia. International journal of remote sensing. 2013 Apr 20;34(8):2730-48.

[6] Kotarba AZ. Regional high-resolution cloud climatology based on MODIS cloud detection data. International Journal of Climatology. 2015 Oct 1.

[7] Irish RR. Landsat 7 automatic cloud cover assessment. InAeroSense 2000 2000 Aug 23 (pp. 348-355). International Society for Optics and Photonics.

[8] Zhu Z, Woodcock CE. Object-based cloud and cloud shadow detection in Landsat imagery. Remote Sensing of Environment. 2012 Mar 15;118:83-94.

[9] Li J, Menzel WP, Yang Z, Frey RA, Ackerman SA. High-spatial-resolution surface and cloud-type classification from MODIS multispectral band measurements. Journal of Applied Meteorology. 2003 Feb;42(2):204-26.

[10] Tian B, Shaikh MA, Azimi-Sadjadi MR, Haar TH, Reinke DL. A study of cloud classification with neural networks using spectral and textural features. IEEE Transactions on Neural Networks. 1999 Jan;10(1):138-51.

[11] Kuo KS, Welch RM, Sengupta SK. Structural and textural characteristics of cirrus clouds observed using high spatial resolution LANDSAT imagery. Journal of Applied Meteorology. 1988 Nov;27(11):1242-60.

[12] Chen DW, Sengupta SK, Welch RM. Cloud field classification based upon high spatial resolution textural features: 2. simplified vector approaches. Journal of Geophysical Research: Atmospheres. 1989 Oct 20;94(D12):14749-65.

[13] Changhui Y, Yuan Y, Minjing M, Menglu Z. Cloud detection method based on feature extraction in remote sensing images. ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 2013 May;1(1):173-7.

[14] Pal M. Ensemble of support vector machines for land cover classification. International journal of remote sensing. 2008 May 20;29(10):3043-9.

[15] Ghimire B, Rogan J, Galiano VR, Panday P, Neeti N. An evaluation of bagging, boosting, and random forests for land-cover classification in Cape Cod, Massachusetts, USA. GIScience & Remote Sensing. 2012 Sep 1;49(5):623-43.

[16] Waske, B., van der Linden, S., Benediktsson, J.A., Rabe, A. and Hostert, P., 2010. Sensitivity of support vector machines to random feature selection in classification of hyperspectral data. IEEE Transactions on Geoscience and Remote Sensing, 48(7), pp.2880-2889.

[17] Pal, M. and Foody, G.M., 2010. Feature selection for classification of hyperspectral data by SVM. IEEE Transactions on Geoscience and Remote Sensing, 48(5), pp.2297-2307.

[18] Huang, M.L., Hung, Y.H., Lee, W.M., Li, R.K. and Jiang, B.R., 2014. SVM-RFE based feature selection and Taguchi parameters optimization for multiclass SVM classifier. The Scientific World Journal, 2014.

[19] Kuncheva LI. Combining pattern classifiers: methods and algorithms. John Wiley & Sons; 2004 Aug 20.

[20] Li X, Wang L, Sung E. AdaBoost with SVM-based component classifiers. Engineering Applications of Artificial Intelligence. 2008 Aug 31;21(5):785-95.

[21] Friedman J, Hastie T, Tibshirani R. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). The annals of statistics. 2000;28(2):337-407.

[22] Warmuth MK, Liao J, Rätsch G. Totally corrective boosting algorithms that maximize the margin. InProceedings of the 23rd international conference on Machine learning 2006 Jun 25 (pp. 1001-1008). ACM.

[23] Breiman L. Random forests. Machine learning. 2001 Oct 1;45(1):5-32.

[24] Gislason PO, Benediktsson JA, Sveinsson JR. Random forests for land cover classification. Pattern Recognition Letters. 2006 Mar 31;27(4):294-300.

[25] Jin J. A Random Forest Based Method for Urban Land Cover Classification using LiDAR Data and Aerial Imagery.

[26] Li F, Clausi DA, Wong A. Comparative study of classification methods for surficial materials in the Umiujalik Lake region using RADARSAT-2 polarimetric, Landsat-7 imagery and DEM data. Canadian Journal of Remote Sensing. 2015 Jan 2;41(1):29-39.

[27] di Bisceglie M, Episcopo R, Galdi C, Ullo SL. Destriping MODIS data using overlapping field-of-view method. IEEE Transactions on Geoscience and Remote Sensing. 2009 Feb;47(2):637-51.

[28] Mather, Paul.M. (2005). Computer Processing of Remotely-Sensed Images: An Introduction, John wiley & Sons, UK.

[29] Hutchison, K.D. and Jackson, J.M., 2003. Cloud detection over desert regions using the 412 nanometer MODIS channel. Geophysical Research Letters, 30(23).

[30] Ackerman S, Strabala K, Menzel P, Frey R, Moeller C, Gumley L. Discriminating clear-sky from cloud with MODIS algorithm theoretical basis document (MOD35. InMODIS Cloud Mask Team, Cooperative Institute for Meteorological Satellite Studies, University of Wisconsin 2010.

[31] Belgiu M, Drăguţ L. Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing. 2016 Apr 30;114:24-31.