



ارزیابی روش انتخاب ویژگی برای شناسایی بهترین باندهای طیفی تصویر ابرطیفی هایپریون (مطالعه موردی: شناسایی دگرسانی‌های معدن مس پورفیری دره زار - کرمان)

امیر سلیمی^{۱*}، منصور ضیائی^۱، علی امیری^۲ و مهدیه حسینجانی‌زاده^۳

(۱) گروه اکتشاف معدن، دانشکده مهندسی معدن، نفت و ژئوفیزیک، دانشگاه صنعتی شاهرود، شاهرود، ایران

(۲) گروه مهندسی کامپیوتر، دانشکده مهندسی، دانشگاه زنجان، زنجان، ایران

(۳) گروه اکولوژی، پژوهشگاه علوم و تکنولوژی پیشرفته و علوم محیطی، دانشگاه تحصیلات تکمیلی صنعتی و فناوری پیشرفته، کرمان، ایران

دریافت مقاله: ۱۳۹۵/۰۷/۳۰، پذیرش: ۱۳۹۵/۱۱/۱۹

چکیده

طبقه‌بندی نظارت شده یکی از بهترین روش‌های استخراج اطلاعات از تصویر است. دقت طبقه‌بندی نظارت شده تصاویر سنجش از دور متأثر از تعداد باندهای طیفی و تعداد نمونه‌های آموزشی است؛ به طوری که اگر تعداد نمونه‌ها از تعداد باندها خیلی کمتر باشد دقت طبقه‌بندی کاهش می‌یابد. وقوع این مشکل در طبقه‌بندی تصاویر ابرطیفی به دلیل داشتن باندهای طیفی بسیار بیشتر در مقایسه با تصاویر چندطیفی، محتمل‌تر است. در چنین شرایطی استفاده از روش‌های انتخاب ویژگی می‌تواند راه‌گشا باشد. این روش‌ها با انتخاب باندهای دارای بیشترین اطلاعات در مورد موضوع مورد بررسی و حذف باندهای نامربوط، زائد و دارای اطلاعات تکراری کارآیی طبقه‌بندی را افزایش می‌دهند. در این پژوهش استفاده از یک روش انتخاب ویژگی مبتنی بر الگوریتم رقابت استعماری فازی به شناسایی تنها ۵ درصد از کل باندهای طیفی یعنی ۹ باند از بین ۱۶۵ باند قابل استفاده هایپریون منجر شد. ۹ باند یادشده بیشترین سهم در شناسایی مناطق دگرسانی را داشته و دقیقاً در بازه‌هایی قرار گرفته‌اند که مناسب برای شناسایی کانی‌های شاخص مناطق دگرسانی هستند که خود بیانگر دقت بالای روش استفاده شده است. همچنین طبقه‌بندی نظارت شده توسط ماشین بردار پشتیبان و با استفاده از ۱۷ نمونه دگرسانی برداشت شده از محدوده نشان می‌دهد که دقت نقشه پراکندگی دگرسانی با استفاده از ۹ باند انتخاب شده در مقایسه با حالتی که کل باندهای هایپریون استفاده شده‌اند، به طور چشمگیری بهبود یافته است.

واژه‌های کلیدی: روش انتخاب ویژگی، مناطق دگرسانی، سنجش از دور ابرطیفی، مس پورفیری دره زار

مقدمه

نیز از این فناوری با اهداف مختلفی بهره‌برداری کرده است؛ به طوری که متخصصان زمین‌شناسی از بدو تولد سنجش از دور، از آن برای تهیه نقشه‌های ناحیه‌ای، تفسیر ساختاری منطقه،

علوم مختلف از تصاویر سنجش از دور برای شناسایی و طبقه‌بندی عوارض سطح زمین استفاده می‌کنند. علم زمین‌شناسی

(Camps-Valls, 2014). یکی از حوزه‌های چالشی مرتبط با داده‌های ابرطیفی، طبقه‌بندی این نوع داده‌هاست. طبقه‌بندی یکی از روش‌های رایج استخراج اطلاعات از تصویر و یا به عبارتی تبدیل داده‌های موجود در تصویر به اطلاعات معنادار است (Oommen, 2008). طبقه‌بندی نظارت‌شده^۴ که بر اساس مجموعه‌ای از داده‌های برچسب‌دار به نام نمونه‌های آموزشی^۵ است، به دلیل حجم زیاد داده‌های ابرطیفی با مشکلات انفجار ابعادی و پدیده هیوز^۶ همراه است. این مشکلات می‌توانند دقت طبقه‌بندی داده‌های ابرطیفی را تحت تأثیر قرار دهند (Plaza et al., 2009; Waske et al., 2009; Camps-Valls et al., 2014).

در حقیقت اساس روش‌های طبقه‌بندی نظارت‌شده، یادگیری قوانین حاکم بر الگوها با استفاده از ویژگی‌های^۷ اندازه‌گیری شده برای نمونه‌های آموزشی است (این ویژگی‌ها در سنجش از دور همان باندهای طیفی هستند). به طور کلی کم یا ناکافی بودن تعداد نمونه‌های آموزشی در مقابل تعداد ویژگی‌ها، به وقوع مشکل انفجار ابعادی منجر می‌شود و در نتیجه آن دقت طبقه‌بندی کاهش می‌یابد و مشکل هیوز رخ می‌دهد که یکی از مشکلات بد حالت^۸ در طبقه‌بندی مجموعه داده‌های بیش ابعاد^۹ است (Plaza et al., 2009; Gheyas and Smith, 2010; Camps-Valls et al., 2014).

رابطه بین تعداد نمونه‌های آموزشی، ابعاد فضای ویژگی^{۱۰} و دقت طبقه‌بندی در شکل ۱ نشان داده شده است (Landgrebe, 2002). چنان‌که مشاهده می‌شود، تغییرات دقت طبقه‌بندی ارتباطی مستقیم با تعداد نمونه‌های آموزشی و ویژگی‌ها دارد؛ اما پس از رسیدن به بیشترین مقدار دقت، ارتباط آن نسبت به فضای ویژگی‌ها معکوس شده و با افزایش تعداد ویژگی‌ها کاهش می‌یابد. این مشکل از آنجا ناشی می‌شود که با ثابت ماندن تعداد نمونه‌های آموزشی در مقابل افزایش تعداد ویژگی‌ها، نمونه‌ها در چنین فضای بزرگی دارای تراکم کمتری بوده (Plaza et

زمین‌شناسی محیطی و پی‌جویی ذخایر معدنی و هیدروکربنی استفاده کرده‌اند (Van der Meer et al., 2012). معمولاً پی‌جویی ذخایر معدنی توسط سنجش از دور با شناسایی سنگ‌ها و کانی‌های شاخصی که مرتبط با ذخیره خاصی هستند، انجام می‌شود که در مقایسه با آنومالی‌های ژئوشیمیایی یا ژئوفیزیکی از آنها با عنوان آنومالی‌های لیتولوژیکی یاد می‌شود (Carranza, 2002). سنگ‌های دگرسان‌شده توسط محلول‌های گرمابی نوعی از آنومالی‌های لیتولوژیکی هستند که از لحاظ موقعیت رخداد، ذخایر را در بر گرفته است و در نتیجه می‌توانند به عنوان راهنمای اکتشافی مورد استفاده قرار گیرند. تهیه نقشه مناطق دگرسان‌شده از طریق بررسی تصاویر سنجش از دور و با هدف اکتشاف ذخایر معدنی به‌وفور و توسط محققان فراوانی انجام شده است (Hosseinjani Zadeh and Tangestani, 2011; Amer et al., 2012; Shahriari et al., 2013; Hosseinjani Zadeh et al., 2014a; Hosseinjani Zadeh et al., 2014b; Shahriari et al., 2015; Salimi et al., 2015). سنجنده‌های سنجش از دور ابرطیفی^۱ در مقایسه با سنجنده‌های چندطیفی^۲ دارای باندهای طیفی بیشتر و در نتیجه از توان تفکیک طیفی^۳ بالاتری برخوردار هستند. این توانایی آنها را به ثبت اطلاعات طیفی دقیق‌تری از عارضه‌های سطح زمین قادر می‌سازد؛ به طوری که می‌تواند اختلافات کوچک طیفی را نیز ثبت کنند (Chang, 2007; Camps-Valls et al., 2014). بنابراین سنجش از دور ابرطیفی توانایی بیشتری در شناسایی کانی‌های مختلف خواهد داشت (Wang and Zheng, 2010) و به همین دلیل در دو دهه اخیر تبدیل به ابزاری مهم برای بررسی سنگ‌ها و کانی‌های تشکیل‌دهنده سطح زمین شده است (Zhang and Peijun, 2014).

تفکیک‌پذیری طیفی بالای داده‌های ابرطیفی امکان بررسی دقیق‌تر عارضه‌های سطح زمین را فراهم می‌کند؛ ولی در مقابل به ظهور چالش‌های محاسباتی جدیدی نیز منجر می‌شود

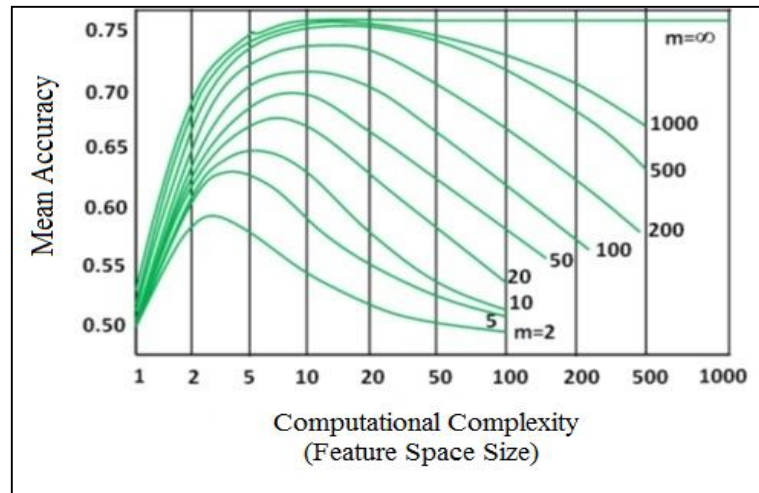
1. Hyperspectral
2. Multispectral
3. Spectral resolution
4. Supervised classification

5. Training samples
6. Curse of dimensionality and Hughes phenomenon
7. Features
8. Ill-posed problem

9. High dimensional data
10. Feature space size
11. Sparse

نمی‌شوند. این مشکل در نهایت به کاهش دقت طبقه‌بندی منجر می‌شود (Waske et al., 2009).

مؤلفه‌های آماری مربوط به داده‌ها به درستی تخمین زده (al., 2009; Camps-Valls et al., 2014) و در نتیجه



شکل ۱. رابطه بین میانگین دقت طبقه‌بندی با ابعاد فضای ویژگی و تعداد نمونه‌های آموزشی ثابت (m) (Landgrebe, 2002)
Fig. 1. Mean accuracy respect to size of feature space and fixed training samples (m) (Landgrebe, 2002)

(Wu, 2013).

به دلیل باندهای طیفی بسیار زیاد، مجموعه داده‌های ابرطیفی را نیز می‌توان یک مجموعه داده بیش ابعاد به‌شمار آورد. داده‌های بیش ابعاد شامل تعداد زیادی از ویژگی‌ها هستند؛ ولی معمولاً بخش زیادی از آنها نامربوط و افزونه^۲ بوده و تنها باعث افزایش اندازه و پیچیدگی فضای ویژگی و همچنین وقوع مشکلات فراوانی مثل هیوز می‌شوند. یکی از راهکارهای موجود برای غلبه بر این مشکل، استفاده از روش‌های کاهش ابعاد فضای ویژگی^۳ است (Alajlan et al., 2012; Camps-Valls et al., 2014). این روش‌ها به‌عنوان گامی مهم در مرحله پیش پردازش داده‌ها، بر روی داده‌های بیش ابعاد اجرا شده است و با انتخاب یا تولید زیرمجموعه کوچک‌تری از ویژگی‌ها موجب افزایش سرعت، دقت و تعمیم‌پذیری طبقه‌بندی می‌شوند (Jimenez and Landgrebe, 1998; Gheyas and Smith, 2010). در حقیقت ویژگی‌های موجود در این مجموعه ویژگی کوچک‌تر، برای استخراج بخش بزرگی از

امروزه در داده‌کاوی، شناسایی الگو، پردازش تصویر، بیوانفورماتیک و بسیاری از حوزه‌های دیگر ممکن است به مسائل بیش ابعادی برخورد شود که در آنها تعداد نمونه‌ها بسیار کم؛ ولی در مقابل تعداد ویژگی‌های توصیف‌گر نمونه‌ها بسیار زیاد است (Yusta, 2009). در عمل و به‌دلایل مختلف، بسیاری از ویژگی‌های مجموعه داده‌های بیش ابعاد، افزونه و نامربوط هستند که فقط باعث بزرگ‌تر شدن اندازه فضای ویژگی می‌شوند. از سویی دیگر تجزیه و تحلیل فضاهای بیش ابعاد و شناسایی الگوهای حاکم بر آنها به‌دلیل پیچیده‌تر بودن با مشکل روبه‌رو می‌شود (Gheyas and Smith, 2010). برای مقابله با این مشکل، روش‌های کاهش ابعاد مورد توجه بسیاری از علوم کاربردی قرار گرفته است. استفاده از این روش‌ها باعث افزایش دقت و سرعت طبقه‌بندی، کاهش پیچیدگی محاسباتی و هزینه زمانی در مرحله آموزش طبقه‌بند، کاهش احتمال رخداد بیش برآزش^۱ و افزایش قدرت تعمیم‌پذیری طبقه‌بندی می‌شود (Yusta, 2009; Gheyas and Smith, 2010; Wang and

1. Over fitting
 2. Irrelevant and redundant
 3. Feature space reduction methods

ویژگی بزرگ نیازمند محاسبات سنگینی است. برای حل این مشکل، روش‌های جستجوگر زیرمجموعه ویژگی‌ها توسعه یافته‌اند که در طی سه مرحله مبادرت به شناسایی ویژگی‌های منتخب می‌کنند. نخست، زیرمجموعه‌ای از ویژگی‌ها را به صورت تصادفی انتخاب کرده، سپس ویژگی‌هایی با وزن بزرگ‌تر از یک حد آستانه را جدا کرده و در پایان از یک تابع ارزیابی برای محاسبه همبستگی بین ویژگی‌ها و شناسایی و حذف ویژگی‌های افزونه استفاده می‌کنند. روش‌های جستجوگر زیرمجموعه ویژگی‌ها در مقایسه با انواع رتبه‌دهی از بازدهی بیشتری برخوردار هستند (Wang and Wu, 2013).

برخلاف روش‌های پالایش، روش‌های پوشاننده از یک مدل یادگیری و یک طبقه‌بند برای شناسایی زیرمجموعه ویژگی‌های منتخب استفاده می‌کنند. شیوه کاری این روش‌ها به این صورت است که ابتدا توسط مدل یادگیری، زیرمجموعه‌های مختلفی از ویژگی‌ها از فضای ویژگی اولیه انتخاب و در ادامه کارآیی آنها توسط یک طبقه‌بند ارزیابی می‌شود. در نهایت زیرمجموعه‌ای با بهترین کارآیی به عنوان زیرمجموعه ویژگی منتخب، انتخاب می‌شود. الگوریتم‌های پوشاننده در مقایسه با انواع پالایش هزینه محاسباتی بیشتری دارند و برای کار با مجموعه داده‌های بیش ابعاد مناسب نیستند؛ ولی از دقت بیشتری برخوردار هستند (Hui-Huang and Cheng-Wei, 2011).

عملکرد روش‌های ترکیبی بر اساس ترکیب روش‌های پالایش و پوشاننده است. این روش‌ها در گام اول با استفاده از یک روش پالایش، تعداد ویژگی‌های مورد نیاز را کاهش می‌دهند، سپس در گام دوم با استفاده از یک روش پوشاننده، بهترین ویژگی‌ها را انتخاب می‌کنند (Bermejo et al., 2012).

از آنجایی که همه باندهای طیفی در تصاویر ابرطیفی برای انجام یک طبقه‌بندی دقیق مورد نیاز نیستند (Bioucas-Dias et al., 2013)، هدف اصلی در این پژوهش استفاده از یک روش انتخاب ویژگی برای پیش پردازش داده‌های ابرطیفی هایپریون و

اطلاعات موجود در داده‌ها کافی خواهند بود (Fukunaga, 1990). روش‌های کاهش ابعاد داده‌ها به دو دسته روش‌های استخراج ویژگی^۱ و انتخاب ویژگی^۲ تقسیم می‌شوند. در نوع اول، فضای ویژگی اولیه دگرگون شده و یک زیرمجموعه جدیدی از ویژگی‌ها تولید می‌شود؛ به طوری که بخش بزرگی از اطلاعات موجود در داده‌ها را پوشش می‌دهد. روش تحلیل مؤلفه‌های اصلی^۳ مثالی از این نوع است. در نوع دوم تعدادی از ویژگی‌ها انتخاب می‌شوند؛ به طوری که این تعداد اندک از ویژگی‌ها حاوی بیشترین اطلاعات موجود در کل ویژگی‌ها هستند (Pal and Foody, 2010). یکی از مزایای نوع انتخاب ویژگی کمک به محققان در شناسایی بهترین ویژگی‌ها در فرآیند طبقه‌بندی است، بنابراین با استفاده از این روش‌ها امکان شناسایی تأثیرگذارترین باندهای طیفی برای طبقه‌بندی مناطق دگرسانی فراهم می‌شود.

روش‌های انتخاب ویژگی به سه دسته روش‌های پالایش، پوشاننده و ترکیبی^۴ طبقه‌بندی می‌شوند. انواع پالایش به هریک از ویژگی‌ها، وزنی متناسب با ویژگی‌های آماری و درجه اهمیت آنها اختصاص می‌دهد و در نهایت با انتخاب ویژگی‌های حائز شرایط مورد نیاز، ویژگی‌های باقی‌مانده را حذف می‌کنند (Wang and Wu, 2013). این روش‌ها که مستقل از الگوریتم‌های یادگیری هستند، خود به دو زیرگروه رتبه‌دهی و جستجوگر زیرمجموعه ویژگی‌ها^۵ تقسیم می‌شوند (Lazar et al., 2012). در نوع رتبه‌دهی با محاسبه همبستگی موجود بین ویژگی‌ها و برجسب طبقات، وزنی به هر طبقه تخصیص یافته و در ادامه ویژگی‌هایی با وزن‌های بیش از حد آستانه، واجد شرایط قرارگیری در بین ویژگی‌های منتخب می‌شوند. عملکرد روش‌های رتبه‌دهی در رابطه با شناسایی ویژگی‌های نامربوط موفقیت‌آمیز است؛ اما در شناسایی ویژگی‌های افزونه دچار مشکل می‌شوند. برای شناسایی باندهای افزونه باید همبستگی بین ویژگی‌ها محاسبه شود که به‌ویژه در رابطه با فضاهای

1. Feature extraction

2. Feature subset selection (FSS)

3. Principal Component Analysis (PCA)

4. Filter, wrapper and hybrid methods

5. Ranking and Subset search

زیادی کانی‌های اکسید آهن دار قرمز تا زردرنگ نیز در منطقه دیده می‌شود (Ranjbar et al., 2001). منطقه دگرسانی تقریباً بیضی‌شکل با طول ۲/۲ و عرض ۱-۰/۷ کیلومتر است؛ به طوری که بخش بزرگی از منطقه توسط دگرسانی‌های پروپلیتیک و فلیک و درصد کمتری نیز با دگرسانی آرژیلیکی پوشیده شده است. همچنین دگرسانی پتاسیک گسترش چندانی در بخش‌های سطحی منطقه ندارد (Geological survey of Iran, 1973).

روش‌ها و داده‌ها

تصویر هایپریون

سنجده فضا برد هایپریون با مشخصاتی بهبود یافته و هزینه کمتر توسط گروه TRW Space and Electronics توسط سازمان فضا نوردی ایالات متحده آمریکا (ناسا) ساخته شد (Chang, 2007). با ارسال سنجده هایپریون توسط EO-1^۲ در نوامبر سال ۲۰۰۰ میلادی، بررسی ابرطیفی سطح زمین از فضا امکان پذیر شد. سنجده هایپریون با داشتن یک تلسکوپ و دو طیف سنج، بازه طیفی ۴۰۰-۲۵۰۰ نانومتری را توسط ۲۴۲ باند طیفی و با دقت فضایی ۳۰^۳ متر پوشش می‌دهد. این باندهای طیفی در مورد بسیاری از کانی‌های سازنده سطح زمین، اطلاعات فراوانی را در اختیار قرار می‌دهند (Liao et al., 2000). داده‌های ارائه شده توسط هایپریون در مقایسه با سنجده‌های هوا برد از کیفیت کمتری برخوردار است؛ اما ارائه اطلاعات با ارزش طیفی توسط آن باعث شده است که کاربردهای متنوعی داشته باشد (Chang, 2007). داده‌های ابرطیفی هوا برد مثل های مپ^۴ و اویریس^۵ که از اوایل دهه هشتاد میلادی در دسترس قرار گرفته‌اند، از توانایی بالایی برای شناسایی کانی‌ها و تشکیلات سنگی برخوردار هستند؛ اما معمولاً دسترسی به این داده‌ها به دلایل مختلف مثل هزینه بالا یا پوشش مکانی کم به سختی امکان پذیر است (Beiranvand Pour and Hashim, 2014; Hosseinjani Zadeh et al., 2014a). بر خلاف داده‌های

انتخاب باندهایی است که بیشترین اطلاعات را برای طبقه بندی مناطق دگرسانی فراهم می‌آورند.

پس از انتخاب باندهای مناسب، طبقه بندی توسط طبقه بند ماشین بردار پشتیبان^۱ و با استفاده از باندهای انتخابی یاد شده، انجام شده است و همچنین برای مقایسه نتایج توسط کل باندهای طیفی تکرار می‌شود. معدن مس پورفیری دره زار در هشت کیلومتری جنوب شرقی معدن سرچشمه به عنوان منطقه مورد بررسی این پژوهش انتخاب شده است.

منطقه مورد بررسی

منطقه مورد بررسی در جنوبی ترین بخش کمر بند ماگمایی ارومیه - دختر قرار دارد. کمر بند یاد شده به سن ژوراسیک و با گسترش تقریبی ۱۸۰۰ کیلومتر از شمال غرب تا جنوب شرق کشور، میزان کانه زایی‌های مهم مس، مولیبدن و طلاست که از جمله این ذخایر می‌توان به معادن سرچشمه، سونگون و میدوک اشاره کرد (شکل ۲-۲A). مجموعه ماگمایی ارومیه - دختر در نتیجه فرورانش صفحه عربی به زیر صفحه ایران مرکزی تشکیل شده و در ادامه در طول دوره‌های اتوسن میانی و الیگوسن تحت فعالیت‌های شدید ماگمایی قرار گرفته است (Alavi, 1980).

محدوده مورد بررسی معدن دره زار در ۸ کیلومتری جنوب شرق معدن سرچشمه قرار گرفته است (شکل ۲-۲B). ذخیره تخمینی معدن دره زار در حدود ۶۷ میلیون تن با عیار ۰/۳۷ درصد است (Hosseinjani Zadeh et al., 2014a). سنگ‌های آتشفشانی رسوبی اتوسن متشکل از آندزیت و تراکی آندزیت میزبان تشکیلات نفوذی دیوریتی تا گرانودیوریتی الیگوسن - میوسن هستند (شکل ۲-۲C). محلول‌های گرمایی موجب تشکیل دگرسانی‌های پتاسیک، فلیک، آرژیلیک و پروپلیتیک در سنگ‌های منطقه شده‌اند (شکل ۳). همچنین در نتیجه فرآیندهای سوپرژن و شست و شوی اسیدی سولفیدها، مقادیر

1. Support Vector Machine (SVM)

2. Earth Observing 1

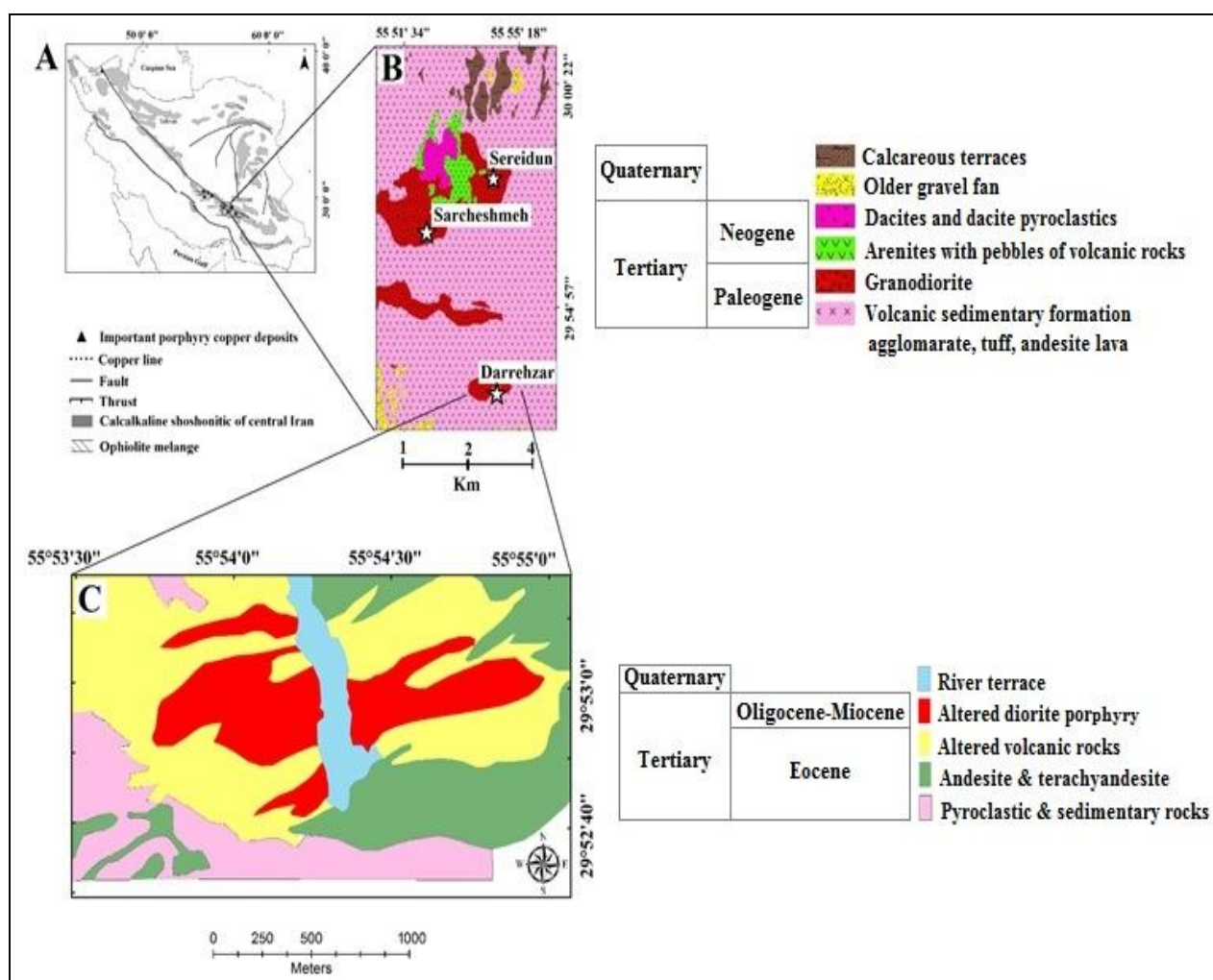
3. Spatial resolution

4. HyMap

5. AVIRIS

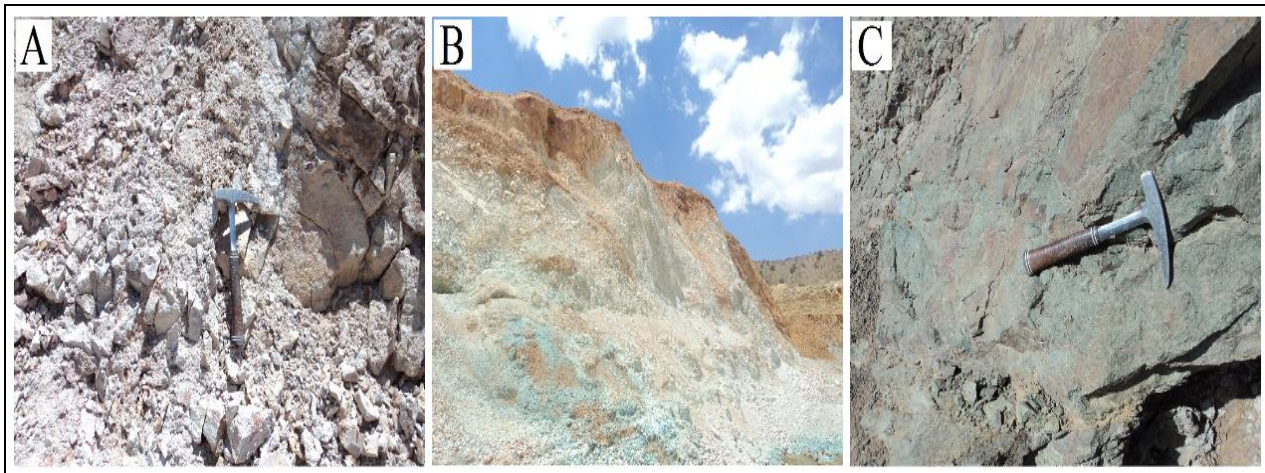
مشخصات مجموعه داده ابرطیفی استفاده شده در این پژوهش که توسط سنجنده هایپریون در ۲۶ ژوئیه ۲۰۰۴ میلادی ثبت شده، در شکل ۴-A نشان داده شده است. همچنین موقعیت جغرافیایی این مجموعه داده نسبت به شهرهای کرمان، رفسنجان و سیرجان در شکل ۴-B قابل مشاهده است. در شکل ۴-C که بخشی از مجموعه داده یادشده را با جزئیات بیشتری نشان می‌دهد، موقعیت معدن دره‌زار نسبت به معدن سرچشمه دیده می‌شود.

هوابرد، دسترسی آسان به داده‌های سنجنده هایپریون، فرصت‌های جدیدی را در اختیار جوامع سنجش از دور قرار داده است تا از اطلاعات طیفی ارزشمند این سنجنده در حوزه‌های مختلف استفاده شود (Petropoulos et al., 2012). بنابراین سنجنده هایپریون در مناطقی که امکان دسترسی به تصاویر هوابرد وجود ندارد، به‌عنوان ابزاری برای انجام بررسی‌های کانی‌شناختی ابرطیفی می‌تواند مورد استفاده قرار گیرد (Kruse, 2003).



شکل ۲. A: موقعیت جغرافیایی محدوده مورد بررسی در ایران، B: نقشه زمین‌شناسی منطقه در مقیاس ناحیه‌ای شامل معدن سرچشمه و دره‌زار برگرفته از حسین‌جانی‌زاده و همکاران (Hosseinjani Zadeh et al., 2014a) و C: نقشه زمین‌شناسی محدوده دره‌زار در مقیاس محلی برگرفته از رنجبار و همکاران (Ranjbar et al., 2001)

Fig. 2. A: Location of study area in Iran, B: Regional scale geological map including Sarcheshmeh and Darrehzar mines (Modified from Hosseinjani Zadeh et al., 2014a), and C: Local scale geological map of Darrehzar area (Modified from Ranjbar et al., 2001)



شکل ۳. دگرسانی‌های معدن دره‌زار: A: آرژیلیک، B: فیلیک و C: پروپیلیتیک

Fig. 3. Alterations of Darrehzar mine A: Argillic, B: Phyllic, and C: Propylitic

نسبت به میانگین اندازه‌گیری شده، نرمالایز می‌کند (Ede, 2004). این الگوریتم اثر مشکلات یادشده را کاهش می‌دهد؛ ولی همچنان مشکلاتی مثل پیکسل‌های غیرنرمال با مقادیر DN^6 منفی در برخی باندها دیده می‌شود. با توجه به قابلیت حل بخش اعظم این مشکلات به صورت بصری، مشکل راه راه شدگی و پیکسل‌های غیرنرمال باقی‌مانده شناسایی و بقیه حذف شدند. با حذف باندهای حاوی پیکسل‌های غیرنرمال به اضافه باندهای واقع شده در بازه طیفی جذبی توسط مولکول‌های بخار آب (باندهای ۱۲۱-۱۳۰ در بازه $1447/2-1356/4$ نانومتر و باندهای ۱۶۵-۱۸۰ در بازه $1800/29-1951/57$ نانومتر)، در نهایت ۱۶۵ باند مطابق با جدول ۱ باقی‌ماند. برای انجام تصحیحات اتمسفری نیز از الگوریتم $FLASH^7$ تعبیه‌شده در نرم‌افزار ENVI استفاده شد.

مجموعه داده‌های آموزشی و آزمایشی^۸

در این پژوهش از دو مجموعه داده با اندازه‌های مختلف و با دو هدف متفاوت استفاده شده است. مجموعه داده اول با تعداد نمونه‌های کمتر که از محدوده مورد بررسی جمع‌آوری شده

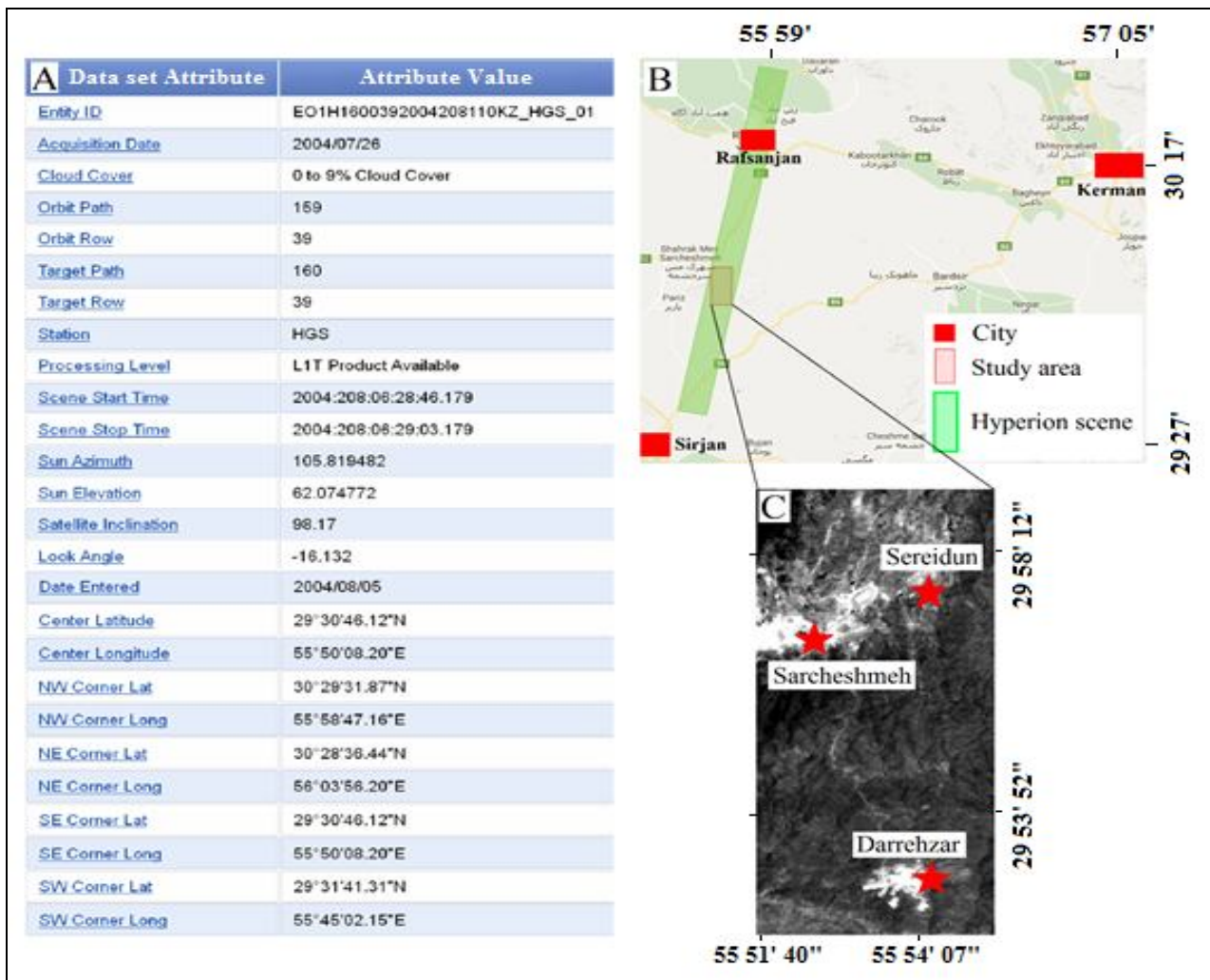
برای افزایش کیفیت و حذف نویز، لازم است مجموعه داده یادشده تحت پیش پردازش‌های لازم قرار گرفته و نسبت به مشکلاتی مثل پیکسل‌های غیرنرمال^۱، راه راه شدگی^۲ و اسمایل^۳ تصحیح شود. سنجنده هایپریون بازتاب امواج الکترومغناطیسی را در ۲۴۲ باند طیفی ثبت می‌کند؛ ولی به دلیل نامطلوب بودن مقادیر مربوط به ۴۴ باند و همچنین وجود ۲ باند افزونه به علت هم‌پوشانی مناطق طیفی VNIR و SWIR، در عمل تنها ۱۹۶ باند از بین ۲۴۲ باند هایپریون باقی‌ماند و پیش پردازش‌های لازم بر روی آنها انجام شد. این ۱۹۶ باند در بازه $925/41-426/82$ محدوده طیفی VNIR و $2395/5-932/64$ محدوده طیفی SWIR قرار دارند. در گام نخست عملیات پیش پردازش، باندهای غیرکالیبره و دارای هم‌پوشانی حذف شدند. با توجه به وجود مشکل راه راه شدگی به‌ویژه در ۱۲ باند ابتدایی منطقه VNIR و بسیاری از باندهای منطقه SWIR، برای کاهش این مشکل از یک الگوریتم بهبود مشکل راه راه شدگی^۴ در نرم‌افزار ENVI^۵ استفاده شد (Ede, 2004; Darmawan, 2006; Hosseinjani Zadeh et al., 2014a). این الگوریتم میانگین هر ردیف را محاسبه و سپس داده‌های آن ردیف را

1. Abnormal pixels
2. Striping
3. Smile
4. De-strip algorithm

5. ENvironment for Visualizing Images
6. Digital Number
7. Fast Line-of sight Atmospheric Analysis of Spectral Hypercubes
8. Test samples

طیفی MTMF^۱ است و همچنین مجموعه داده نخست در ادامه معرفی شده‌اند.

است، برای طبقه‌بندی مناطق دگرسانی و مجموعه داده دوم با تعداد نمونه‌های بیشتر برای کاهش ابعاد مجموعه داده هایپریون استفاده خواهند شد. این مجموعه داده که خروجی روش تجزیه



شکل ۴. A: مشخصات مجموعه داده ثبت شده توسط سنجنده هایپریون، B: موقعیت جغرافیایی تصویر هایپریون نسبت به شهرهای بزرگ و C: بخشی از تصویر هایپریون و موقعیت معادن دره‌زار و سرچشمه بر روی آن

Fig. 4. A: Specifications of Hyperion dataset, B: Location of Hyperion scene regard to main cities, and C: A part of the Hyperion image and location of Darrehzar and Sarcheshmeh mines

آموزش و ارزیابی طبقه‌بند ماشین بردار پشتیبان از ۱۷ نمونه سنگی برداشت‌شده از مناطق دگرسان‌شده معدن دره‌زار استفاده شد. نمونه‌های یادشده توسط دستگاه طیف‌سنج ASD^۲ در

روش‌های طبقه‌بندی نظارت‌شده برای یادگیری قوانین حاکم بر الگوها و همچنین برای ارزیابی میزان این یادگیری به ترتیب نیازمند مجموعه داده‌های آموزشی و آزمایشی هستند. برای

1. Mixture Tuned Matched Filtering
2. Analytical Spectral Device FieldSpc3 (ASD)

آزمایشگاه طیفی سنجی دانشگاه تحصیلات تکمیلی صنعتی و فناوری پیشرفته کرمان آنالیز و نتایج به دست آمده در کنار نظر کارشناسی برای مشخص کردن نوع دگرسانی نمونه‌ها استفاده شدند (شکل ۵-۸).

جدول ۱. ۱۶۵ باند قابل استفاده تصویر هایپریون (شامل معدن دره‌زار) پس از انجام پیش پردازش‌های لازم

Table 1. 165 Useable spectral bands of Hyperion (Including Darrehzar mine) after pre-processing

Electromagnetic range	Bands	Wavelength (nm)
VNIR	8-57	426-925
	79-93	932-1073
	95-98	1094-1124
	100-115	1144-1295
	117-120	1316-1346
SWIR	131-164	1457-1790
	181-189	1961-2042
	191-202	2062-2173
	204-224	2193-2395

زیرپیکسلی است و فراوانی هر یک از کانی‌ها را در هر پیکسل تخمین می‌زند، بنابراین برای اطمینان بیشتر و داشتن خطای کمتر در ساخت این مجموعه داده، پیکسل‌هایی از تصویر انتخاب شد که فراوانی یکی از کانی‌ها بیش از ۵۰ درصد باشد. در نهایت پیکسل‌های انتخابی در طبقه این کانی قرار گرفتند و برچسب‌گذاری شدند. در شکل ۵-B و جدول ۲ موقعیت و تعداد نمونه‌های این مجموعه داده دیده می‌شود.

روش انتخاب ویژگی

در این پژوهش برای انتخاب مجموعه ویژگی‌های منتخب از یک الگوریتم ترکیبی استفاده می‌شود. در این الگوریتم در مرحله پالایش، ویژگی‌ها وزن‌دهی شده، سپس در مرحله پوشاننده و با استفاده از الگوریتم رقابت استعماری فازی^۱ و الگوریتم IWSSr^۲، بهترین زیرمجموعه ویژگی‌ها انتخاب

کاهش ابعاد مجموعه داده هایپریون با استفاده از الگوریتم به‌کاررفته در این پژوهش (که در ادامه معرفی شده است) نیازمند استفاده از یک مجموعه داده با تعداد نمونه‌های فراوان است. از آنجایی که تهیه چنین مجموعه داده‌ای از منطقه مورد بررسی نیازمند صرف هزینه و زمان زیادی است، سعی شد که این مجموعه به‌صورت مصنوعی تهیه شود. بدین منظور، نخست روش تجزیه طیفی MTMF برای طبقه‌بندی دگرسانی مناطق سرچشمه، دره‌زار و سریدون استفاده و سپس خروجی به‌دست آمده به‌عنوان مجموعه داده آموزشی منظور شد. این روش پیش از این به‌عنوان روش زیرپیکسلی دقیق برای شناسایی کانی‌های دگرسانی در منطقه مورد بررسی این پژوهش توسط حسین‌جانی‌زاده و همکاران (Hosseinjani Zadeh et al., 2014a) استفاده شده و با نتایج مطلوبی نیز همراه بوده است. البته باید اشاره شود از آنجایی که روش MTMF، روشی

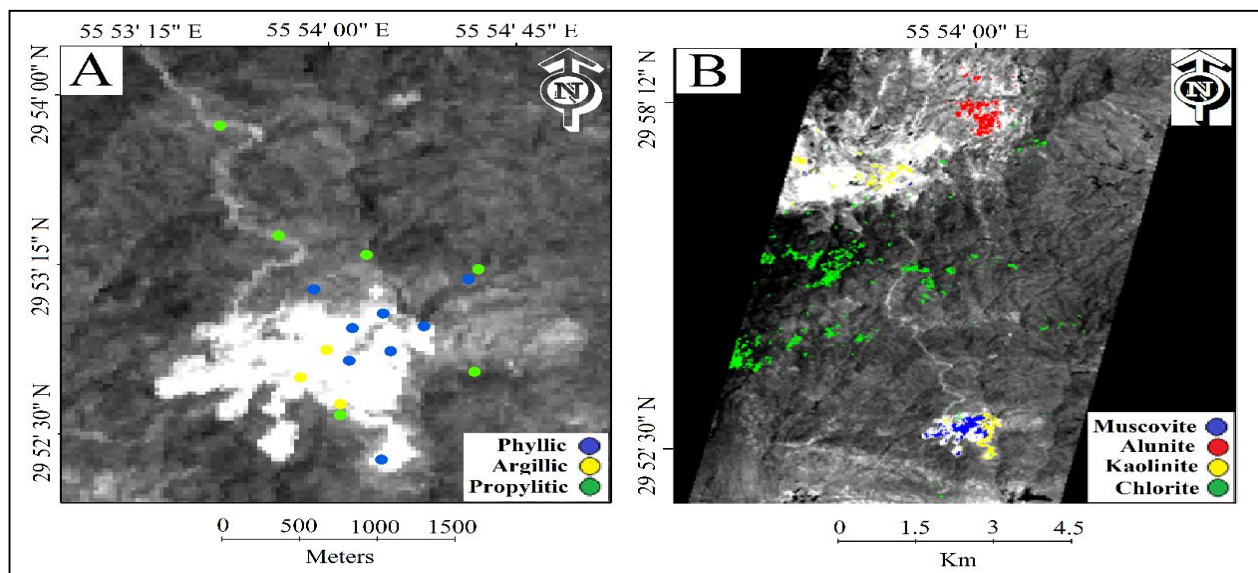
1. Fuzzy Imperialist Competitive Algorithm (FICA)

2. Incremental Wrapper Subset Selection with replacement

مؤثر در شناسایی مناطق دگرسانی مورد استفاده قرار گرفتند. سپس برای ارزیابی عملکرد باندهای منتخب به‌دست آمده در مرحله قبل، طبقه‌بندی مناطق دگرسانی با استفاده از کل باندهای طیفی هایپریون و همچنین باندهای انتخابی، انجام شده و در نهایت نتایج به‌دست آمده، مقایسه می‌شوند. همان‌گونه که در شکل نیز دیده می‌شود طبقه‌بند ماشین بردار پشتیبان و مجموعه داده معرفی شده در شکل ۵-A برای اجرای مرحله طبقه‌بندی استفاده خواهند شد. برای انجام هر دو مرحله یادشده و کسب خروجی‌های مورد نیاز از نرم افزار متلب^۳ استفاده شد.

می‌شود. این روش علاوه بر دستیابی به زیرمجموعه فشرده‌تری از ویژگی‌ها از دقت بیشتری نیز در مقایسه با روش‌های مشابه برخوردار است. برای دستیابی به جزئیات بیشتری از این الگوریتم به مرادخانی و همکاران (Moradkhani et al., 2015) مراجعه شود.

در شکل ۶، مراحل انجام کار طی دو مرحله مجزا نشان داده شده است. در گام نخست، الگوریتم انتخاب ویژگی یادشده و همچنین مجموعه داده معرفی شده در شکل ۵-B، برای کاهش ابعاد طیفی تصویر هایپریون و با هدف انتخاب باندهای طیفی



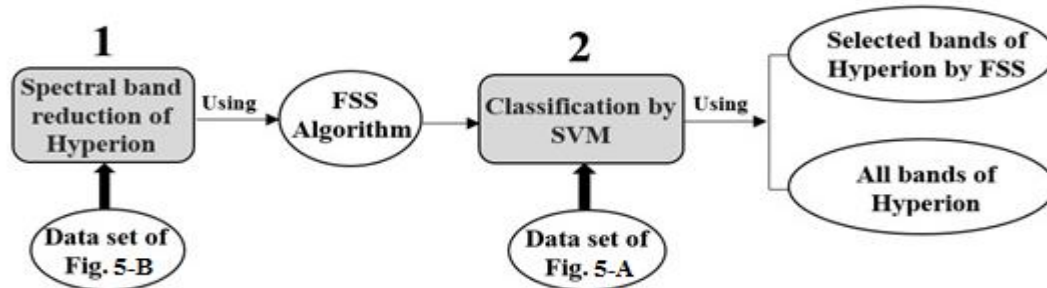
شکل ۵. مجموعه داده‌های استفاده شده برای A: آموزش طبقه‌بند SVM (نمونه‌برداری شده از معدن دره‌زار) و B: کاهش ابعاد تصویر هایپریون

Fig. 5. Utilized datasets for A: SVM training (Sampled from Darrehzar mine), and B: Reduction of Hyperion size

جدول ۲. مشخصات مجموعه داده استفاده شده برای کاهش ابعاد مجموعه داده هایپریون

Table 2. Characteristics of used dataset for feature reduction of Hyperion

Class (mineral)	Size (pixels)	Related alteration	Color
Muscovite	379	Phyllic	Blue
Alunite	557	Advanced argillic	Red
Kaolinite	504	Argillic	Yellow
Chlorite	1700	Propylitic	Green



شکل ۶. مراحل انجام کار ۱. کاهش تعداد باندهای تصویر هایپریون با استفاده از الگوریتم انتخاب ویژگی و مجموعه داده معرفی شده در شکل ۵-B، ۲. طبقه‌بندی مناطق دگرسانی توسط ماشین بردار پشتیبان و مجموعه داده معرفی شده در شکل ۵-A. به منظور مقایسه، طبقه‌بندی مناطق دگرسانی معدن دره زار، یکبار با استفاده از کل باندهای هایپریون و بار دیگر توسط باندهای انتخابی توسط الگوریتم انتخاب ویژگی انجام شد.

Fig. 6. Flowchart of process: 1. Reduction of Hyperion bands using FSS and dataset of Fig. 5B, 2. Classification of alteration zones using SVM and dataset of Fig. 5A. For comparison, the alteration zones of Darrehzar mine were classified two times, first using all spectral bands of Hyperion and then using selected spectral bands of Hyperion resulted from FSS algorithm.

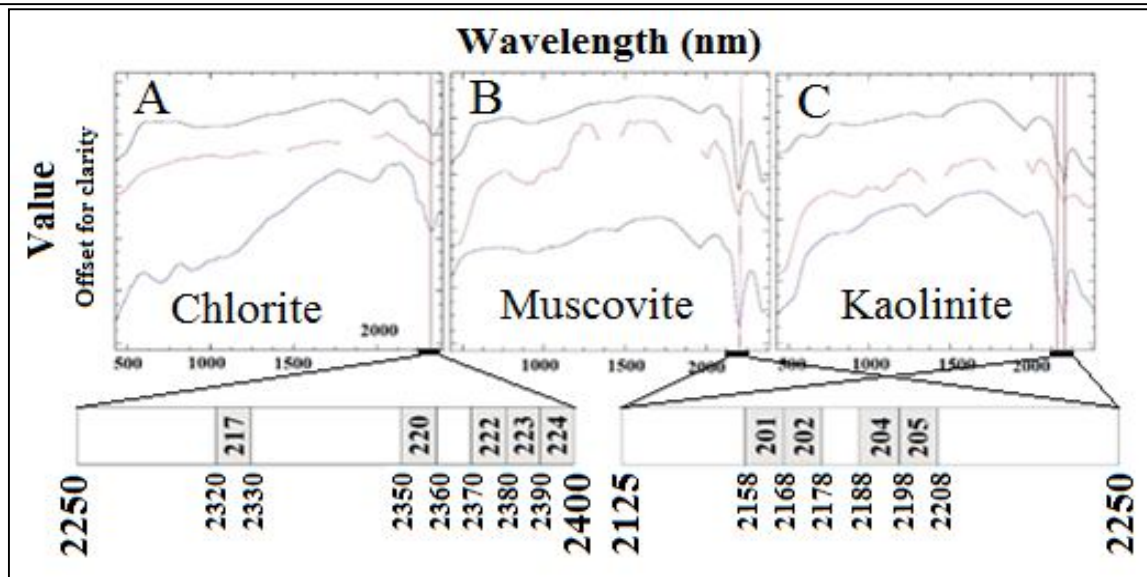
بحث و بررسی

چنان‌که گفته شد، تعداد باندهای فراوان داده‌های ابرطیفی امتیازی مهم برای این نوع داده‌هاست و باعث افزایش قابلیت آنها در شناسایی دقیق‌تر عارضه‌های سطح زمین می‌شود؛ اما احتمال آنکه تعداد زیادی از این باندها افزونه و نامربوط باشند نیز خیلی زیاد است. این باندها نه تنها مشارکتی در شناسایی عوارض ندارند؛ بلکه به پیچیده شدن فرآیند طبقه‌بندی نیز منجر می‌شوند. بنابراین قبل از اجرای طبقه‌بندی و برای کم کردن تعداد باندها، استفاده از روش‌های کاهش ابعاد فضای ویژگی می‌تواند سودمند باشد. هدف این روش‌ها کم کردن اندازه و پیچیدگی مجموعه داده اولیه است؛ به گونه‌ای که فرآیند طبقه‌بندی با سهولت و دقت بیشتری انجام شود.

در این پژوهش استفاده از الگوریتم انتخاب ویژگی معرفی شده در بخش قبل، به شناسایی ۹ باند طیفی تاثیرگذار در شناسایی کانی‌های دگرسانی از بین ۱۶۵ باند تصویر هایپریون منجر شد. ۹ باند انتخاب شده متعلق به بازه طیفی ۲۰۰۰-۲۵۰۰ نانومتر هستند که در منابع علمی مختلف به نقش مهم این بازه طیفی در شناسایی کانی‌های رسی مناطق دگرسانی اشاره شده است (Gersman et al., 2008; Hosseinjani Zadeh et al., 2014c). بر اساس شکل ۷ که موقعیت این باندها را در بازه طیفی یاد شده با جزئیات بیشتری نشان می‌دهد، باندهای شماره

۲۰۱، ۲۰۲، ۲۰۴ و ۲۰۵ متعلق به بازه ۲۱۲۵-۲۲۵۰ نانومتر و باندهای ۲۱۷، ۲۲۰، ۲۲۲، ۲۲۳ و ۲۲۴ متعلق به بازه ۲۲۵۰-۲۴۰۰ هستند. با توجه به نمودارهای بالای تصویر مشاهده می‌شود که باندهای یاد شده دقیقاً در بازه‌هایی قرار گرفته‌اند که کانی‌های شاخص مناطق دگرسانی دارای ویژگی‌های طیفی متمایز و شاخصی (منطبق با خطوط عمودی) هستند. یعنی مسکویت و کائولینیت جذب طیفی شاخصی در بازه ۲۱۲۵-۲۲۵۰ نانومتر و به خصوص در باندهای ۲۰۱ و ۲۰۵ دارند و کلریت نیز در بازه ۲۴۰۰-۲۲۵۰ نانومتر و در باند ۲۱۸ دارای جذب طیفی شاخصی است (Hosseinjani Zadeh et al., 2014a). بنابراین تطابق باندهای انتخابی توسط الگوریتم انتخاب ویژگی با بازه‌های طیفی یاد شده می‌تواند بیانگر عملکرد مطلوب روش انتخاب ویژگی باشد.

البته همان‌گونه که دیده می‌شود الگوریتم انتخاب ویژگی استفاده شده چندین باند دیگر که در مجاورت باندهای ۲۰۱، ۲۰۵ و ۲۱۸ هستند را نیز به عنوان باندهای مؤثر در شناسایی مناطق دگرسانی معرفی کرده است. علت این موضوع می‌تواند ناشی از وجود همبستگی و هم‌پوشانی طیفی بالا بین باندهای طیفی مجاور در تصویر هایپریون باشد (Bioucas-Dias et al., 2013).

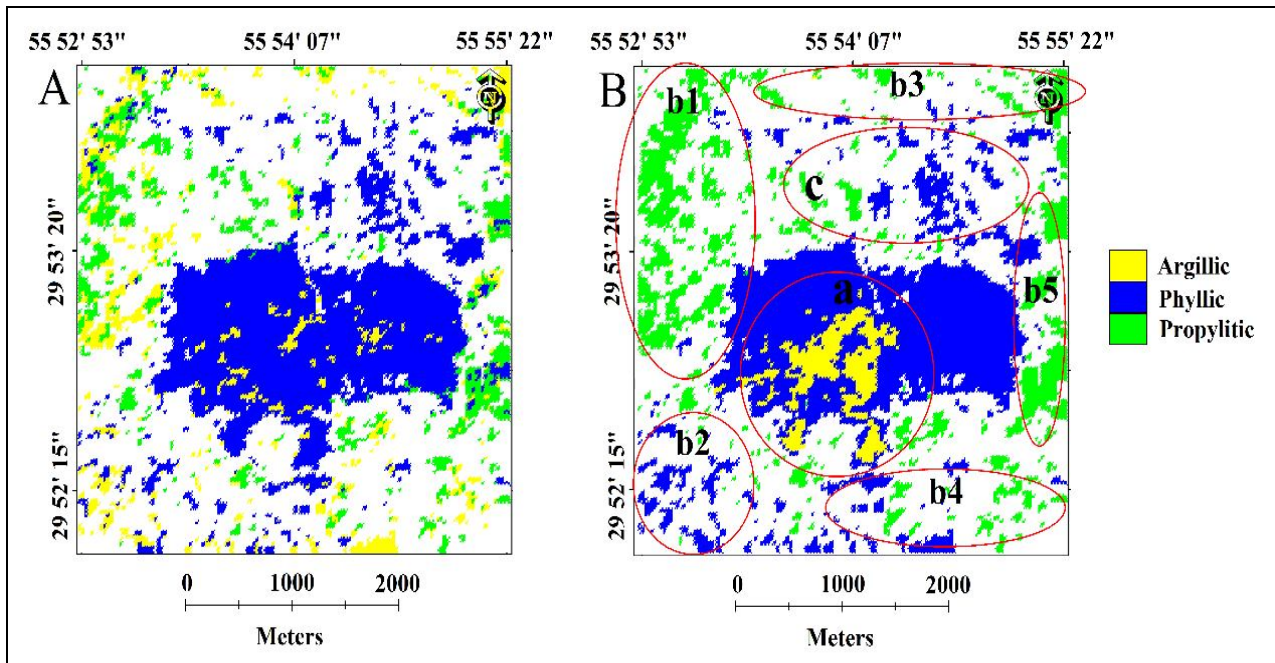


شکل ۷. بالا: ویژگی‌های طیفی کانی‌های شاخص مناطق دگرسانی A: کلریت، B: مسکویت و C: کائولینیت. بازه‌ای که جذب شاخص طیفی برای هر کانی رخ می‌دهد، با خط سیاه‌رنگ افقی مشخص شده است. برگرفته از حسین‌جانی‌زاده و همکاران (Hosseinjani Zadeh et al., 2014a). پایین: ۹ باند طیفی به‌دست آمده توسط الگوریتم انتخاب ویژگی

Fig. 7. Top: Spectral characteristic of indicator minerals of alteration zones A: Chlorite, B: Muscovite, and C: Kaolinite. A range of diagnostic absorption features of each mineral has been highlighted by a horizontal black line (Modified from Hosseinjani Zadeh et al., 2014a), down: 9 spectral bands outputted from feature subset selection algorithm

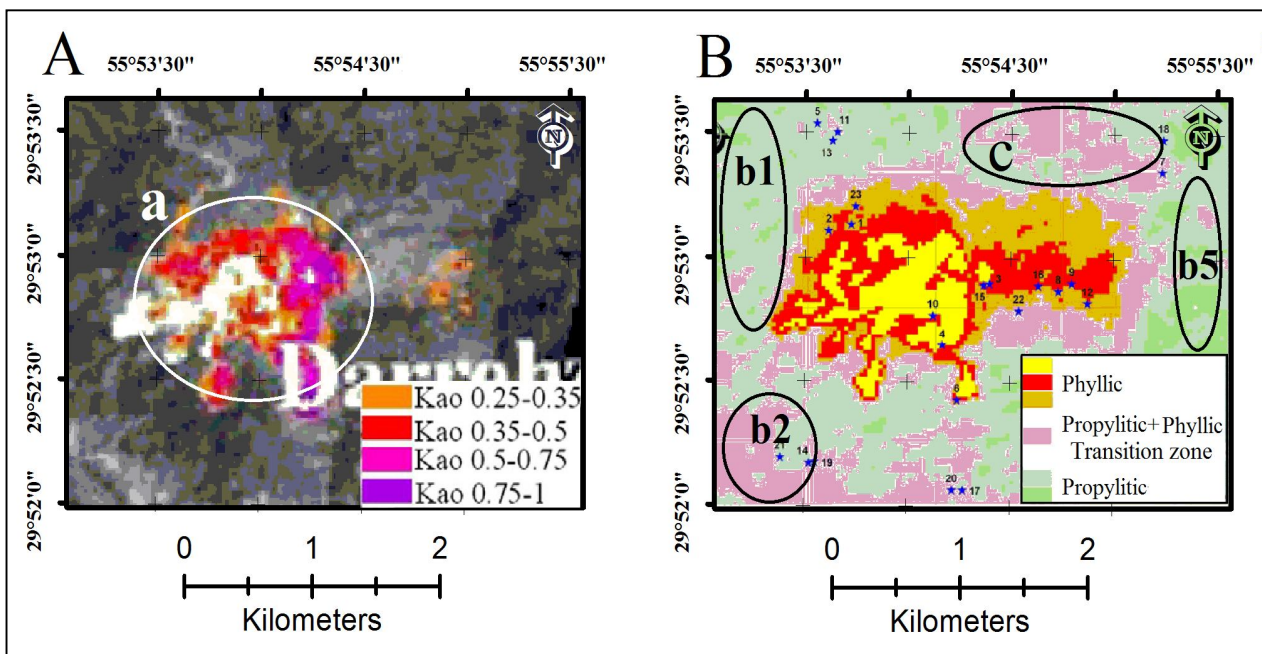
چنان‌که دیده می‌شود در نقشه به‌دست آمده از ۹ باند منتخب، بخش مرکزی محدوده (منطقه a) که تقریباً منطبق بر پیت معدن دره زار نیز هست به‌عنوان منطقه آرژیلیکی طبقه‌بندی شده است؛ در حالی‌که بخش غالب این محدوده در نقشه حاصله از ۱۶۵ باند طیفی به‌عنوان منطقه فیلیک شناسایی شده بود. افزایش دقت نتایج طبقه‌بندی در نتیجه استفاده از ۹ باند منتخب، در دیگر مناطق مشخص شده در شکل ۸-B (مناطق b1, b2, b3, b4, b5) دیده می‌شود. برای مثال بر اساس پژوهش‌های شهریار و همکاران (Shahriari et al., 2013) (شکل ۹-B)، مناطق b1 و b5 توسط دگرسانی پروپیلیتییک پوشیده شده‌اند؛ اما بخش‌هایی از این مناطق در شکل ۸-A به اشتباه به‌عنوان دگرسانی‌های آرژیلیک و فیلیک طبقه‌بندی شده‌اند که در شکل ۸-B اصلاح شده‌اند. همچنین بخش بزرگی از منطقه b2 نیز در شکل ۸-A به اشتباه به‌عنوان دگرسانی آرژیلیکی شناسایی شده‌اند. این خطا باز هم در شکل ۸-B با شناسایی درست این مناطق تحت عنوان دگرسانی فیلیک تصحیح شده است.

پس از انتخاب باندهای مؤثر در شناسایی کانی‌های دگرسانی، طبقه‌بندی مناطق دگرسانی منطقه توسط طبقه‌بند ماشین بردار پشتیبان، یک‌بار با کلیه باندهای طیفی تصویر هایپریون (۱۶۵ باند) و بار دیگر با ۹ باند طیفی منتخب انجام شد که نتایج به‌دست آمده در شکل ۸ نشان داده شده است. بر اساس شکل ۸-A که نقشه دگرسانی به‌دست آمده با استفاده از کلیه باندهای هایپریون را نشان می‌دهد، مشاهده می‌شود که دگرسانی آرژیلیکی (ناحیه زرد رنگ) در بخش وسیعی از منطقه مورد بررسی حضور دارد. این درحالی است که بر اساس مطالب گفته‌شده در بخش زمین‌شناسی منطقه، نمونه‌های برداشت‌شده از منطقه طی بازدیدهای صحرائی و همچنین پژوهش حسین‌جانی‌زاده و همکاران (Hosseinjani Zadeh et al., 2014a) (شکل ۹-A)، دگرسانی آرژیلیکی گسترش زیادی در منطقه نداشته و تقریباً محدود به بخش‌های مرکزی منطقه است. این خطای طبقه‌بندی تا حدود بسیار زیادی در نقشه دگرسانی به‌دست آمده از ۹ باند طیفی رفع شده است (شکل ۸-B).



شکل ۸. نقشه دگرسانی طبقه‌بندی شده معدن دره‌زار به‌دست آمده از طبقه‌بند ماشین بردار پشتیبان با استفاده از A: کل ۱۶۵ باند طیفی هایپریون و B: ۹ باند طیفی معرفی شده توسط الگوریتم انتخاب ویژگی

Fig. 8. Classified alteration maps of Darrehzar mine resulted from SVM classifier using A: All 165 spectral bands of Hyperion, and B: 9 selected spectral bands of FSS algorithm



شکل ۹. نقشه دگرسانی معدن دره‌زار A: برگرفته از حسین‌جانی‌زاده و همکاران (Hosseinjani Zadeh et al., 2014a) و B: برگرفته از شهریاری و همکاران (Shahriari et al., 2013)

Fig. 9. Alteration maps of Darrehzar mine A: Modified from Hosseinjani Zadeh et al. (2014a), and B: Modified from Shahriari et al. (2013)

یادشده باعث شد دقت نتایج افزایش یابد و خطاهای طبقه‌بندی توسط ۱۶۵ باند طیفی تا حدود زیادی رفع شود. در حقیقت کاهش تعداد باندهای طیفی از میزان پیچیدگی فضای طیفی کاست و امکان شناسایی قوانین حاکم بر الگوی مناطق دگرسانی را با دقت بیشتری فراهم کرد؛ به طوری که در نتیجه آن تفکیک دگرسانی‌ها با دقت خیلی بیشتری انجام شد. بنابراین می‌توان گفت که استفاده از روش‌های انتخاب ویژگی در مرحله پیش پردازش داده‌های ابرطیفی به دستیابی نتایج قابل اعتمادتری منجر خواهد شد. این کار به‌ویژه برای آن دسته از داده‌های ابرطیفی که مشکل عدم قطعیت و نویز دارند (مثل هایپریون) و دارای باندهای افزونه فراوانی هستند، ضروری‌تر است

همچنین شناسایی مؤثرترین باندهای طیفی برای تفکیک مناطق دگرسانی، دیگر مزیت الگوریتم انتخاب ویژگی استفاده شده است که آن را از روش‌های استخراج ویژگی (برای مثال روش آنالیز مؤلفه‌های اصلی) متمایز می‌کند. این موضوع از آن جهت حائز اهمیت است که پژوهشگران با آن دسته از باندهای طیفی که بیشترین اطلاعات را درباره مناطق دگرسانی فراهم می‌کنند آشنا شده و مستقیماً می‌توانند این باندها را برای پژوهش‌های مشابه مورد استفاده قرار دهند. در حالی که به‌عنوان مثال در نتیجه استفاده از روش آنالیز مؤلفه‌های اصلی امکان کاهش ابعاد فضای طیفی وجود دارد؛ ولی امکان شناخت بهترین باندهای طیفی وجود ندارد.

قدردانی

نویسندگان از پرسنل شرکت ملی صنایع مس ایران به‌ویژه آقای مهندس خسروجریدی در بخش تحقیق و توسعه بابت حمایت‌های بی‌دریغ و از آقای مهندس سلاجقه کارشناس محترم معدن دره‌زار برای راهنمایی‌های علمی راهگشا، قدردانی می‌کنند.

البته باید اشاره شود که بر اساس شکل ۹-B، این منطقه و همچنین منطقه C که به‌عنوان مناطق فیلیک طبقه‌بندی شده‌اند جزو دگرسانی فیلیک قوی نبوده و منطقه‌گذاری بین دگرسانی‌های فیلیک و پروپلیتیک است (Shahriari et al., 2013). یعنی پیکسل‌های تصویر در این مناطق تلفیقی از دو دگرسانی یادشده هستند. به بیان دیگر، مناطق آبی رنگ در بخش‌های b2 و C، هر دو به‌عنوان دگرسانی فیلیک شناسایی شده‌اند؛ اما حقیقت این است که دگرسانی فیلیک در این مناطق از شدت کمتری نسبت به مناطق مرکزی‌تر برخوردار است. این خطا از آنجا به‌وجود آمده است که ماشین بردار پشتیبان یک طبقه‌بند پیکسل پایه بوده و قادر به انجام طبقه‌بندی در مقیاس زیر پیکسلی نیست؛ بنابراین توانایی تخمین فراوانی دگرسانی‌های مختلف در پیکسل‌های ترکیبی را نداشته و تنها پیکسل‌ها را بر اساس شدت دگرسانی و با نام دگرسانی غالب برجسب زده و طبقه‌بندی می‌کند (Salimi et al., 2016).

نتیجه‌گیری

طبقه‌بندی نظارت‌شده تصاویر سنجنش از دور ابرطیفی به‌دلیل فراوانی تعداد باندهای طیفی این نوع تصاویر تحت‌تأثیر مشکل هیوز قرار گرفته است و از دقت آن کاسته می‌شود. این مشکل در صورتی که تعداد نمونه‌های آموزشی برای آموزش طبقه‌بند در مقایسه با تعداد باندهای طیفی خیلی کمتر باشد، می‌تواند تشدید شود. برای غلبه بر این مشکل روش‌های کاهش ابعاد فضای ویژگی مورد استفاده قرار می‌گیرند.

در این پژوهش از یک روش انتخاب ویژگی مبتنی بر الگوریتم رقابت استعماری فازی برای شناسایی باندهای طیفی تأثیرگذار در شناسایی کانی‌های دگرسانی استفاده شد که در نتیجه آن فقط در حدود ۵٪ از کل باندهای طیفی هایپریون (۹ باند) به‌عنوان باندهای منتخب شناسایی شدند. طبقه‌بندی توسط ۹ باند

References

- Alajlan, N., Bazi, Y., Melgani, F. and Yager, R., 2012. Fusion of supervised and unsupervised learning for improved classification of hyperspectral images. *Journal of Information Sciences*, 217: 39–55.
- Alavi, M., 1980. Tectono-stratigraphic evolution of the Zagros side of Iran. *Geology*, 8(3): 144–149.
- Amer, R., Kusky, T. and El Mezayen, A., 2012. Remote sensing detection of gold related alteration zones of Um Rus Area, Central Eastern Desert of Egypt. *Advances in Space Research*, 49(1): 121–134.
- Beiranvand Pour, A. and Hashim, M., 2014. ASTER, ALI and Hyperion sensors data for lithological mapping and ore minerals exploration. *SpringerPlus*, 3(1): 1–19.
- Bermejo, P., de la Ossa, L., Gámez, J.A. and Puerta, J.M., 2012. Fast wrapper feature subset selection in high-dimensional datasets by means of filter re-ranking. *Knowledge-based Systems*, 25(1): 35–44.
- Bioucas-Dias, J.M., Plaza, A., Camps-Valls, G., Scheunders, P., Nasrabadi, N.M. and Chanussot, J., 2013. Hyperspectral Remote Sensing Data Analysis and Future Challenges. *IEEE Geoscience and Remote Sensing Magazine*, 1(2): 6–36.
- Camps-Valls, G., Tuia, D., Bruzzone, L. and Benediktsson, J., 2014. Advances in Hyperspectral Image Classification. *IEEE Signal Processing Magazine*, 31(1): 45–54.
- Carranza, E.J.M., 2002. Geologically-Constrained Mineral Potential Mapping. Ph.D. Thesis, Delft University of Technology, Delft, Netherlands, 480 pp.
- Chang, C., 2007. Hyperspectral Data Exploitation, Theory and Applications. John Wiley & Sons, New Jersey, 430 pp.
- Darmawan, A., 2006. Mapping soil mineral using Hyperion imagery in relation to the level of structural damage in the Bam earthquake. M.Sc. Thesis, the University of Melbourne, Melbourne, Australia, 86 pp.
- Ede, R.V., 2004. Destriping and Geometric Correction of an ASTER Level 1A Image. Utrecht University, Utrecht, 36 pp.
- Fukunaga, K., 1990. Statistical Pattern Recognition. Academic Press, San Diego, 592 pp.
- Geological survey of Iran, 1973. Exploration for Ore deposits in Kerman region. Ministry of Economy Geological Survey of Iran, Tehran, Report Yu/53.
- Gersman, R., Ben-Dor, E., Beyth, M., Avigad, D., Abraha, M. and Kibreba, A., 2008. Mapping of hydrothermal altered rocks by the EO-1 Hyperion sensor, northern Danakil, Eritrea. *International Journal of Remote Sensing*, 29(13): 3911–3936.
- Gheyas, A. and Smith, L.S., 2010. Feature subset selection in large dimensionality domains. *Pattern Recognition*, 43(1): 5–13.
- Hosseinjani Zadeh, M. and Tangestani, M.H., 2011. Mapping alteration minerals using sub-pixel unmixing of ASTER data in the Sarduiyeh area, southeastern Kerman Iran. *International Journal of Digital Earth*, 4(6): 487–504.
- Hosseinjani Zadeh, M., Tangestani, M.H., Velasco Roldan, F. and Yusta, I., 2014a. Sub-pixel mineral mapping of a porphyry copper belt using EO-1 Hyperion data. *Advances in Space Research*, 53(3): 440–451.
- Hosseinjani Zadeh, M., Tangestani, M.H., Velasco Roldan, F. and Yusta, I., 2014b. Mineral exploration and alteration zone mapping using mixture tuned matched filtering approach on ASTER data at the central part of Dehaj-Sarduiyeh copper belt, SE Iran. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(1): 284–289.
- Hosseinjani Zadeh, M., Tangestani, M.H., Velasco Roldan, F. and Yusta, I., 2014c.

- Spectral characteristics of minerals in alteration zones associated with porphyry copper deposits in the middle part of Kerman copper belt, SE Iran., *SE Iran. Ore Geology Reviews*, 62: 191-198.
- Hui-Huang, H. and Cheng-Wei, H., 2011. Hybrid feature selection by combining filters and wrappers. *Expert Systems with Applications*, 38(7): 8144–8150.
- Jimenez, L.O. and Landgrebe, D.A., 1998. Supervised classification in high-dimensional space: geometrical, statistical, and asymptotical properties of multivariate data. *IEEE Transactions on Systems, Man, and Cybernetics*, 28(1): 39–54.
- Kruse, F.A., 2003. Mineral Mapping with AVIRIS and EO-1 Hyperion. 12th JPL Airborne Geoscience Workshop, Pasadena, California.
- Landgrebe, D.A., 2002. Hyperspectral Image Data Analysis. *IEEE Signal processing Magazine*, 19(1): 17–28.
- Lazar, C., Taminau, J., Meganck, S., Steenhoff, D., Coletta, A., Molter, C., de Schaet-zen, V., Duque, R., Bersini, H. and Nowe, A., 2012. A survey on filter techniques for feature selection in gene expression microarray analysis. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 9(4): 1106–1119.
- Liao, L., Jarecke, P., Gleichauf, D. and Hedman, T., 2000. Performance and characterization of the Hyperion imaging spectrometer instrument. *International Symposium on Optical Science and Technology*, San Diego, California, USA.
- Moradkhani, M., Amiri, A., Javaherian, M. and Safari, H., 2015. A hybrid algorithm for feature subset selection in high-dimensional data sets using FICA and IWSSr algorithm. *Applied Soft Computing*, 35: 123–135.
- Oommen, T., 2008. An objective analysis of Support Vector Machine based classification for remote sensing. *Mathematical Geosciences*, 40(4): 409–424.
- 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): 2297–2307.
- Petropoulos, G.P., Kalaitzidis, C. and Prasad Vadrevu, K., 2012. Support vector machines and object-based classification for obtaining land-use/cover cartography from Hyperion hyperspectral imagery. *Computers & Geosciences*, 41: 99–107.
- Plaza, A., Benediktsson, J.A., Boardman, J.W., Brazile, J., Bruzzone, L., Camps-Valls, G., Chanussot, J., Fauvel, M., Gamba, P., Gualtieri, A., Marconcini, M., Tilton, J.C. and Trianni, G., 2009. Recent advances in techniques for hyperspectral image processing. *Remote Sensing of Environment*, 113: 110–122.
- Ranjbar, H., Hassanzadeh, H. and Torabi, M., 2001. Integration and analysis of airborne geophysical data of the Darrehzar area, Kerman Province, Iran, using principal component analysis. *Journal of Applied Geophysics*, 48(1): 33–41.
- Salimi, A., Ziaii, M., Amiri, A. and Hosseinjani Zadeh, M., 2016. Sub-pixel classification of hydrothermal alteration zones using a kernel-based method and hyperspectral data; A case study of Sarcheshmeh Porphyry Copper Mine and surrounding area, Kerman, Iran. *Journal of Mining and Environment*, 8(4): 555-565.
- Salimi, A., Ziaii, M., Hosseinjani Zadeh, M., Amiri, A. and Karimpouli, S., 2015. High Performance of the Support Vector Machine in Classifying Hyperspectral Data Using a Limited Dataset. *International Journal of Mining and Geo-Engineering*, 49(2): 253–268.
- Shahriari, H., Honarmand, M. and Ranjbar, H., 2015. Comparison of multi-temporal ASTER images for hydrothermal alteration mapping using a fractal-aided SAM method. *International Journal of Remote Sensing*, 36(5): 1271–1289.

- Shahriari, H., Ranjbar, H. and Honarmand, M., 2013. Image Segmentation for Hydrothermal Alteration Mapping Using PCA and Concentration-Area Fractal Model. *Natural Resources Research*, 22(3): 191–206.
- Van der Meer, F., Van der Werff, H., Van Ruitenbeek, F., Hecker, C., Bakker, W., Noomen, M., Van der Meijde, M., Carranza, E., Boudewijn de Smeth, J. and Woldai, T., 2012. Multi- and hyperspectral geologic remote sensing: A review. *International Journal of Applied Earth Observation and Geoinformation*, 14(1): 112–128.
- Wang, J. and Wu, L., 2013. Maximum weight and minimum redundancy: A novel framework for feature subset selection. *Pattern Recognition*, 46(6): 1616–1627.
- Wang, Z.H. and Zheng, C.Y., 2010. Rocks/Minerals Information Extraction from EO-1 Hyperion Data Base on SVM. *International Conference on Intelligent Computation Technology and Automation*, Changsha, China.
- Waske, B., Benediktsson, J.A., Arnason, K. and Sveinsson, J.R., 2009. Mapping of hyperspectral AVIRIS data using machine-learning algorithms. *Canadian Journal of Remote Sensing*, 35(1): 106–116.
- Yusta, S.C., 2009. Different metaheuristic strategies to solve the feature selection problem. *Pattern Recognition Letter*, 30(5): 525–534.
- Zhang, X. and Peijun, L., 2014. Lithological mapping from hyperspectral data by improved use of spectral angle mapper. *International Journal of Applied Earth Observation and Geoinformation*, 31: 95–109.



Evaluation of a Feature Subset Selection method to find informative spectral bands of Hyperion hyperspectral data for hydrothermal alteration mapping: A case study from the Darrehzar porphyry copper mine, Kerman, Iran

Amir Salimi^{1*}, Mansour Ziaii¹, Ali Amiri² and Mahdiah Hosseinjani Zadeh³

1) School of Mining Engineering, Petroleum and Geophysic, Shahrood University of Technology, Shahrood, Iran

2) Computer Engineering Group, Department of Engineering, University of Zanjan, Zanjan, Iran

3) Department of Ecology, Research Institute of Sciences and High Technology and Environmental Sciences, Graduate University of Advanced Technology, Kerman, Iran

Submitted: Oct. 21, 2016

Accepted: Feb. 7, 2017

Keywords: *Feature subset selection method, Alteration zones, Hyperspectral remote sensing, Darrehzar porphyry type copper deposit*

Introduction

In the regional prospecting of ore minerals, geologists usually utilize remote sensing images for hydrothermal alteration mineral mapping as a kind of lithological anomaly, which may be linked to mineral deposits (Carranza, 2002).

Compared to the multispectral remote sensing images, composed of few spectral bands, the hyperspectral data prepare much more spectral details of the surface materials in many bands. These high spectral resolution images provide subtle spectral data for identifying similar materials of the surface (Camps-Valls et al., 2014). This ability could greatly promote the potential of hyperspectral based mineral mapping (Wang and Zheng, 2010). As in the two last decades, hyperspectral remote sensing has been an important tool for studying earth's minerals and rocks (Zhang and Peijun, 2014).

Although, the high number of spectral bands is an important advantage for hyperspectral images, many of those bands are usually irrelevant and redundant and, therefore, cause just the size and complexity of the band space to be increased. This complexity can lead to an ill-posed problem in supervised classification, namely the curse of

dimensionality and the Hughes phenomenon, which negatively affect the accuracy of the classification (Camps-Valls, 2014).

Feature reduction methods can be applied to overcome these problems and to eliminate those spectral bands in the classification of hyperspectral images that provide no further useful information. These methods produce an efficient subset of features (spectral bands in remote sensing field) from the original feature space. The decrease in complexity obtained as a result of the feature space reduction can increase the ability of classifiers to efficiently capture the classification rules. Consequently, the speed, generalization, and predictive classification accuracy are increased (Gheyas and Smith, 2010; Camps-Valls et al., 2014).

This study is aimed at evaluation and management of the curse of dimensionality risk in hyper spectral data classification by means of a feature reduction method. The method is utilized to select more informative spectral bands of Hyperion hyperspectral data, which are more effective for the classification of hydrothermal alteration zones. The well-known study area here is the Darrehzar

*Corresponding authors Email: a.salimi@shahroodut.ac.ir

porphyry copper mine located 8 km from the southeast of the giant Sarcheshmeh mine.

Materials and methods

1. Hyperion data

The Hyperion hyperspectral image with 242 spectral bands acquired on July 26, 2004 was available and it was used in this study.

2. Train and test datasets

Two datasets were utilized. The first dataset that resulted from the Mixture Tuned Matched Filtering (MTMF) method was applied to feed the feature reduction method and the second dataset containing 17 rock samples collected from the study area was used to carry out the classification by SVM.

3. The feature reduction method

In this study, we applied a hybrid Feature Subset Selection (FSS) method to reduce the number of spectral bands of Hyperion data. Extensive details may be found in Moradkhani et al. (2015).

Discussion and results

The Feature Subset Selection (FSS) algorithm was applied to reduce the size of the spectral bands of Hyperion data. The implementation of this algorithm resulted in the selection of 9 bands among all 165 spectral bands (i.e. 5% of all useable spectral bands of Hyperion) as the more influential bands for the identification of clay minerals. These bands belong to the two spectral ranges, 2125-2250 nm and 2250-2400 nm, respectively. On the other hand, it is believed that the Short-Wave Infrared (SWIR) electromagnetic range (2000-2500 nm) is an important spectral range for distinguishing clay minerals of the hydrothermal alteration systems (Hosseinjani Zadeh et al., 2014). This implies that two ranges introduced by FSS were accurately selected, because both of them coincide with the SWIR range. Clearly speaking, bands 201, 202, 204, and 205 in the range of 2125-2250 nm are used for muskovit, kaolinit and alunite enhancement. Moreover, bands 217, 220, 222, 223, and 224 in the 2250-2400 nm are appropriate for chlorite classification.

A comparison between the maps of SVM based classification of the alteration zones using 9 (selected by feature selection method) and 165 (all useable bands of Hyperion data) spectral bands confirmed a significant improvement in the output results when 9 more informative bands are utilized for classification instead of all 165 bands. In fact, the classification based on 9 selected bands is comparable and even more effective than the full band classification. This is because the decrease in spectral bands makes SVM learn the rules of classification more accurately.

Reference

- Camps-Valls, G., Tuia, D., Bruzzone, L. and Benediktsson, J., 2014. Advances in Hyperspectral Image Classification. *IEEE Signal Processing Magazine*, 31(1): 45–54.
- Carranza, E.J.M., 2002. Geologically-Constrained Mineral Potential Mapping. Ph.D. Thesis, Delft University of Technology, Delft, Netherlands, 480 pp.
- Gheyas, A. and Smith, L.S., 2010. Feature subset selection in large dimensionality domains. *Pattern Recognition*, 43(1): 5-13.
- Hosseinjani Zadeh, M., Tangestani, M.H., Velasco Roldan, F. and Yusta. I., 2014. Spectral characteristics of minerals in alteration zones associated with porphyry copper deposits in the middle part of Kerman copper belt, SE Iran., SE Iran. *Ore Geology Reviews*, 62: 191-198.
- Moradkhani, M., Amiri, A., Javaherian, M. and Safari, H., 2015. A hybrid algorithm for feature subset selection in high-dimensional data sets using FICA and IWSSr algorithm. *Applied Soft Computing*, 35: 123–135.
- Wang, Z.H. and Zheng, C.Y., 2010. Rocks/Minerals Information Extraction from EO-1 Hyperion Data Base on SVM. *International Conference on Intelligent Computation Technology and Automation*, Changsha, China.
- Zhang, X. and Peijun, L., 2014. Lithological mapping from hyperspectral data by improved use of spectral angle mapper. *International Journal of Applied Earth Observation and Geoinformation*, 31: 95–109.