

استخراج ویژگی نظارت شده غیر پارامتریک برای طبقه‌بندی تصاویر ابرطیفی با نمونه‌ی آموزشی محدود

مریم ایمانی^۱، حسن قاسمیان^۲

چکیده

استخراج ویژگی در تحلیل و طبقه‌بندی تصاویر ابرطیفی، اهمیت ویژه‌ای دارد، چراکه علاوه بر بهبود طبقه‌بندی، سبب کاهش پیچیدگی محاسباتی خواهد شد. روش‌های استخراج ویژگی نظارت شده‌ای مثل تحلیل تمیز خطی (LDA) به دلیل مشکل منفرد بودن ماتریس پراکنندگی درون کلاسی، دارای کارایی خوبی در نمونه‌های آموزشی محدود نیستند. به علاوه تعداد ویژگی‌های استخراج شده توسط آن‌ها حداکثر برابر تعداد کلاس‌ها منهای یک است. استخراج ویژگی وزن‌دار غیر پارامتریک (NWFE) این مشکلات را حل کرده، ولی به شدت دارای پیچیدگی محاسباتی بالایی است. در این مقاله، یک روش استخراج ویژگی نظارت شده غیر پارامتریک پیشنهاد شده است که علاوه بر حل مشکلات ذکر شده در روش‌های استخراج ویژگی LDA و NWFE، دارای کارایی بهتری نسبت به آن‌ها در طبقه‌بندی تصاویر ابرطیفی می‌باشد. روش استخراج ویژگی پیشنهادی با سه روش استخراج ویژگی نظارت شده معمول مقایسه شده و نتایج آزمایش‌ها بر روی سه داده‌ی ابرطیفی واقعی، نشان‌دهنده‌ی کارایی مناسب روش پیشنهادی می‌باشد.

کلیدواژه

استخراج ویژگی، طبقه‌بندی، تصاویر ابرطیفی، نمونه‌ی آموزشی محدود

مقدمه

دو زیرمجموعه روش‌های نظارت شده (که با استفاده از نمونه‌های آموزشی کار می‌کنند) و روش‌های بدون نظارت (روش‌هایی که بدون استفاده از نمونه‌های آموزشی کار می‌کنند) تقسیم می‌شوند. البته می‌توان دسته‌ی سوم نیمه نظارت شده را نیز به آن‌ها اضافه کرد که یا از ترکیب روش‌های نظارت شده و بدون نظارت حاصل می‌شود و یا از نمونه‌های آموزشی نیمه‌برچسب‌دار علاوه بر نمونه‌های آموزشی اصلی برای کاهش ویژگی استفاده می‌کند.

فرآیند انتخاب ویژگی به معنای انتخاب زیرمجموعه‌ی مناسبی از ویژگی‌های اولیه و دور ریختن مابقی ویژگی‌هاست. انتخاب موثرترین زیرمجموعه از ویژگی‌ها بر طبق یک تابع معیار و یک الگوریتم جستجو انجام می‌پذیرد [۵]-[۳]. تمرکز اصلی ما در این مقاله، استخراج ویژگی می‌باشد. استخراج ویژگی می‌تواند با استفاده از نمونه‌های آموزشی و به صورت نظارت شده انجام شود، مانند روش معمول و کلاسیک تحلیل تمیز خطی (LDA) [۶] و یا می‌تواند به صورت بدون نظارت (بدون استفاده از نمونه‌های آموزشی) انجام گیرد. معروف‌ترین روش

پیشرفت‌های اخیر حاصل شده در سیستم‌های سنجش از دور، اندازه‌گیری تابش در تعداد باندهای طیفی بسیار زیادی را فراهم می‌کند. مثالی از این تکنولوژی، سنجنده‌ی طیف‌سنج تصویربرداری مرئی-مادون قرمز هوابرد (AVIRIS) است که داده‌ی تصویری با ۲۲۰ باند تولید می‌کند [۱]. بعد افزوده شده این‌گونه تصاویر نسبت به تصاویر چندطیفی، اطلاعات زیادی را برای ما فراهم می‌کند. ولی در عین حال استفاده از روش‌های معمول برای تحلیل تصاویر ابرطیفی کارساز نیست و این موضوع چالش‌برانگیزی برای تحلیل داده‌های ابرطیفی ایجاد می‌کند. کاهش ویژگی داده‌ی ابرطیفی فواید زیادی در تحلیل و طبقه‌بندی تصاویر ابرطیفی دارد که از آن جمله می‌توان به موارد زیر اشاره کرد: ۱- درک و فهم بهتر داده ۲- کاهش پیچیدگی محاسباتی و نیاز به حافظه‌ی کمتر ۳- افزایش دقت طبقه‌بندی [۲]. کاهش ویژگی به دو روش کلی انتخاب و استخراج ویژگی انجام می‌شود. هر کدام از این دو گروه شامل

^۱ دانشجوی دکتری برق، دانشگاه تربیت مدرس

^۲ استاد دانشکده برق، دانشگاه تربیت مدرس، ghassemi@modares.ac.ir

^۳ Airborne Visible/Infrared Imaging Spectrometer

شده است، به دلیل منفرد شدن ماتریس پراکندگی درون-کلاسی، دارای کارایی خوبی در نمونه‌های آموزشی محدود نخواهد بود. روش استخراج ویژگی وزن‌دار غیرپارامتریک (NWFE^۹) که در [۱۴] پیشنهاد شده، دارای کارایی بهتری نسبت به LDA می‌باشد، ولی هزینه‌ی محاسباتی سنگینی دارد. یک روش استخراج ویژگی نظارت شده‌ی دیگر نزدیک‌ترین نسبت دوتایی (DNP^{۱۰}) است که می‌تواند به عنوان بسطی از LDA و NDA در نظر گرفته شود [۱۵]. ساختار به کار گرفته شده در DNP، اثر هم‌پوشانی میان توزیع کلاس‌ها را کاهش می‌دهد و از یک روش تنظیم جدید برای مقابله با منفرد شدن ماتریس پراکندگی درون-کلاسی استفاده می‌کند. روش استخراج ویژگی نظارت شده‌ی دیگر که اخیراً پیشنهاد شده، عامل‌بندی ماتریس غیرمنفی تبدیل متعامد با حفظ همسایگی (NPOPNMF^{۱۱}) می‌باشد [۱۶]. برای به دست آوردن ماتریس تبدیل NPOPNMF از دو عبارت در حل مسئله‌ی بهینه‌سازی آن استفاده می‌شود که یک عبارت مربوط به عامل‌بندی ماتریس غیر منفی تبدیل و عبارت دیگر مربوط به حفظ همسایگی محلی می‌باشد. هر دو روش‌های DNP و NPOPNMF در سایز کوچک نمونه‌ی آموزشی دارای کارایی نزدیک به NWFE می‌باشند و از آن‌جایی که هدف ما در این مقاله، استخراج ویژگی با استفاده از نمونه‌های آموزشی محدود می‌باشد، از بین روش‌های NWFE، DNP و NPOPNMF تنها روش معروف NWFE برای مقایسه با روش پیشنهادی انتخاب گردیده است. علاوه بر NWFE روش پیشنهادی ما با روش‌های پرکاربرد LDA و GDA نیز مقایسه شده است.

در این مقاله یک روش نسبتاً ساده و کارا پیشنهاد شده که در مقایسه با چند روش استخراج ویژگی پرکاربرد بیان شده، در شرایطی که تعداد نمونه‌های آموزشی محدود است، از نظر دقت طبقه‌بندی برتری دارد. به علاوه، روش پیشنهادی مشکلاتی از قبیل منفرد شدن^{۱۲} ماتریس پراکندگی درون-کلاسی، تعداد محدود ویژگی استخراج شده و زمان طولانی محاسباتی را ندارد. روش پیشنهادی با چندین روش معمول استخراج

بدون نظارت استخراج ویژگی، تحلیل مولفه‌های اصلی (PCA^۵) می‌باشد [۶]. دسته اول اشاره شده (روش‌های با نظارت) به دلیل در نظر گرفتن جدایی‌پذیری کلاس‌ها، مناسب برای کاربردهای طبقه‌بندی‌اند. اما دسته‌ی دوم (روش‌های بدون نظارت) تمیز بین کلاس‌ها را در نظر نگرفته و بر اساس معیارهای دیگری از جمله حداقل کردن مربع خطای بازسازی سیگنال، مانند PCA، عمل می‌کنند و در نتیجه ممکن است برای کاربردهای طبقه‌بندی مفید واقع نشوند. هم‌چنین روش‌های نظارت شده، خود به دو دسته روش‌های پارامتریک و غیرپارامتریک، تقسیم‌بندی می‌شوند. علاوه بر این دسته‌بندی، می‌توان یک گروه‌بندی دیگر برای روش‌های استخراج ویژگی در نظر گرفت: روش‌های خطی و غیرخطی. روش‌های غیرخطی معمولاً از تابع هسته^۶ برای بسط مدل خطی به مدل غیرخطی استفاده می‌کنند. با استفاده از یک تابع هسته‌ی مناسب برای تبدیل داده، جدایی‌پذیری کلاس‌ها در فضای کرنل می‌تواند افزایش یابد. از جمله این روش‌ها می‌توان به تحلیل تمیز تعمیم‌یافته (GDA^۷) اشاره کرد که بسط غیرخطی روش LDA است [۷]. امروز نسبتاً کاملی از روش‌های مختلف انتخاب و استخراج ویژگی در [۸] و تعدادی از روش‌های استخراج ویژگی در [۱۱] - [۹] ارائه شده‌اند. نویسندگان در [۱۲] برای استخراج ویژگی نظارت شده از مفهوم اطلاعات متقابل استفاده کرده‌اند. از آن‌جایی که تخمین اطلاعات متقابل، نسبت به آمارگان مرتبه‌ی اول و دوم، به نمونه‌ی آموزشی بسیار بیشتری احتیاج دارد، در نتیجه، این روش با استفاده از سایز کوچک نمونه‌ی آموزشی، قابل اجرا نخواهد بود. روش استخراج ویژگی نظارت شده‌ی دیگر با استفاده از شبکه‌های عصبی در [۱۳] پیشنهاد شده است که این روش هم به دلیل نیاز شبکه عصبی به نمونه‌ی آموزشی زیاد برای یادگیری، با استفاده از نمونه‌ی آموزشی محدود کارا نمی‌باشد.

روش‌های نظارت شده‌ای مثل LDA و GDA به دلیل محدودیت رتبه‌ی^۸ ماتریس پراکندگی بین-کلاسی، قادر به استخراج حداکثر تعداد کلاس‌ها منهای یک ویژگی هستند. به علاوه، روش LDA، که یک روش پرکاربرد و معمول نظارت

^۹ nonparametric weighted feature extraction

^{۱۰} Double Nearest Proportion

^{۱۱} Neighborhood Preserving Orthogonal Projection Nonnegative Matrix

Factorization

^{۱۲} singularity

^۵ Principal component analysis

^۶ Kernel function

^۷ generalized discriminant analysis

^۸ rank

$$S_b = \sum_{k=1}^c n_k (\mathbf{m}_k - \mathbf{m})(\mathbf{m}_k - \mathbf{m})^T \quad (2)$$

$$S_w = \sum_{k=1}^c \left(\sum_{i=1}^{n_k} (\mathbf{x}_{i,k} - \mathbf{m}_k)(\mathbf{x}_{i,k} - \mathbf{m}_k)^T \right) \quad (3)$$

n_k ، \mathbf{m}_k ، \mathbf{m} و $\mathbf{x}_{i,k}$ به ترتیب تعداد نمونه‌های آموزشی کلاس k ام، میانگین کلاس k ام، میانگین کل نمونه‌های آموزشی و نمونه‌ی k ام در کلاس k ام هستند. همچنین تعداد کلاس‌ها برابر C می‌باشد. ویژگی‌های استخراج شده در روش LDA با حداکثر کردن معیار $tr(S_w^{-1}S_b)$ به دست می‌آیند. $tr(A)$ به معنای مجموع عناصر روی قطر اصلی ماتریس A می‌باشد.

سه ایراد اساسی روش LDA عبارت‌اند از:

(الف) LDA زمانی خوب کار می‌کند که کلاس‌ها دارای توزیع نرمال باشند. اگر توزیع کلاس‌ها غیر نرمال یا چند مده باشد، کارایی روش LDA چندان خوب نخواهد بود.

(ب) چون رتبه ماتریس پراکندگی بین-کلاسی برابر تعداد کلاس‌ها منهای یک می‌باشد، LDA حداکثر قادر به استخراج $c - 1$ ویژگی است.

(ج) با وجود نمونه‌های آموزشی محدود، ماتریس پراکندگی درون-کلاسی، منفرد^{۱۳} شده و در این حالت، LDA کارایی ضعیفی خواهد داشت.

تحلیل تمییز تعمیم یافته (GDA)

از طریق نگاشت داده از فضای ورودی به فضای ویژگی با بعد بالاتر، می‌توان روش‌های استخراج ویژگی خطی را به روش‌های غیرخطی بسط داد. از توابع هسته برای محاسبه‌ی ضرب داخلی در فضای بعد بالاتر استفاده می‌شود. GDA بسط غیرخطی LDA است [۷]. با استفاده از یک نگاشت Φ می‌توان نمونه‌ی $\mathbf{x} \in \mathbb{R}^d$ را به فضای بعد بالاتر \mathbb{H} منتقل کرد، $\Phi(\mathbf{x}) \in \mathbb{H}$ در فضای جدید، برای به دست آوردن ماتریس تبدیل برای استخراج n ($n < d$) ویژگی و به دست آوردن $\mathbf{z} = \mathbf{W}^T \mathbf{x}$ ، $\mathbf{z} \in \mathbb{R}^n$ تابع هزینه‌ی زیر باید بهینه شود:

ویژگی بر روی سه داده‌ی واقعی ابرطیفی ایندیانا، سالیاناس و مرکز پاپوا پیاده‌سازی و مقایسه شده است. ادامه مقاله به شکل زیر سازمان یافته است: در قسمت بعدی، تعدادی از روش‌های استخراج ویژگی پرکاربرد که روش پیشنهادی خود را با آن مقایسه کرده‌ایم، بیان شده است. سپس به توصیف روش پیشنهادی خود پرداخته‌ایم. در ادامه، نتایج آزمایش‌ها مورد بحث قرار گرفته‌اند و در نهایت، نتایج به دست آمده به اختصار بیان شده‌اند.

چند روش استخراج ویژگی معمول و پرکاربرد

در این قسمت، تعدادی از روش‌های پرکاربرد استخراج ویژگی را بیان خواهیم کرد. از آنجایی که بحث اصلی ما در این رساله، روش‌های استخراج ویژگی مناسب برای کاربردهای طبقه‌بندی می‌باشد، در نتیجه بر روی روش‌های نظارت شده استخراج ویژگی تمرکز کرده و به بیان و بررسی آن‌ها می‌پردازیم. تعدادی از روش‌های بیان شده در این قسمت، در بخش بعدی با روش پیشنهادی مطرح شده در این مقاله، با یکدیگر مقایسه خواهند شد.

تحلیل تمییز خطی (LDA)

یک روش نظارت شده‌ی پارامتریک است که از معیار جدایی‌پذیری کلاس‌ها برای استخراج ویژگی استفاده می‌کند [۶]. روش LDA، نسبت ماتریس کواریانس بین-کلاسی را به ماتریس کواریانس درون-کلاسی بهینه می‌کند. فرض که بعد داده‌ی اولیه d و بعد فضای تبدیل یافته n باشد، $\{\mathbf{x}_i\}_{i=1}^N$ و $\{\mathbf{z}_i\}_{i=1}^N$ به ترتیب داده بعد بالا (داده‌ی اولیه) و داده بعد پایین (داده‌ی تبدیل شده) و N تعداد کل نمونه‌های داده است. نگاشت داده از \mathbb{R}^d به \mathbb{R}^n با $\mathbf{z}_i \in \mathbb{R}^n$ با یک تبدیل خطی $\mathbf{z}_i = \mathbf{W}^T \mathbf{x}_i$ انجام می‌شود. بردار \mathbf{w} که نشان‌دهنده‌ی بردار تشکیل دهنده‌ی ماتریس تبدیل $\mathbf{W}_{d \times n} = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n)$ می‌باشد، طبق مسئله‌ی بهینه‌سازی زیر به دست می‌آید:

$$\mathbf{w} = \arg \max_{\mathbf{w}} \frac{\mathbf{w}^T S_b \mathbf{w}}{\mathbf{w}^T S_w \mathbf{w}} \quad (1)$$

ماتریس‌های پراکندگی درون-کلاسی S_w و بین کلاسی S_b به شکل زیر تعریف می‌شوند:

^{۱۳} singular

$$S_w = \sum_{i=1}^c p_i \sum_{l=1}^{n_i} \frac{\lambda_l^{(i,i)}}{n_i} (x_l^{(i)} - M_i(x_l^{(i)})) (x_l^{(i)} - M_i(x_l^{(i)}))^T \quad (8)$$

وزن ماتریس پراکندگی $\lambda_l^{(i,j)}$ و میانگین وزن دار $x_l^{(i)}$ در کلاس j ، $M_j(x_l^{(i)})$ ، به صورت زیر تعریف می‌شوند:

$$\lambda_l^{(i,j)} = \frac{\text{dist}(x_l^{(i)}, M_j(x_l^{(i)}))^{-1}}{\sum_{t=1}^{n_i} \text{dist}(x_t^{(i)}, M_j(x_t^{(i)}))^{-1}} \quad (9)$$

$$M_j(x_l^{(i)}) = \sum_{k=1}^{n_i} w_{lk}^{(i,j)} x_k^{(j)}, \quad w_{lk}^{(i,j)} = \frac{\text{dist}(x_l^{(i)}, x_k^{(j)})^{-1}}{\sum_{t=1}^{n_i} \text{dist}(x_l^{(i)}, x_t^{(j)})^{-1}} \quad (10)$$

برای جلوگیری از منفرد شدن ماتریس پراکندگی درون-کلاسی، NWFE از روش تنظیم^{۱۴} به شکل زیر استفاده می‌کند:

$$S_w = 0.5S_w + 0.5\text{diag}(S_w) \quad (11)$$

که $\text{diag}(A)$ به معنای درایه‌های قطری ماتریس A است.

عیب اساسی روش NWFE، زمان محاسباتی بسیار زیاد آن است که با افزایش n_i هم به طور قابل ملاحظه‌ای افزایش می‌یابد.

نزدیک‌ترین نسبت دوتایی (DNP)

روش استخراج ویژگی نظارت شده DNP، اثر هم‌پوشانی میان توزیع کلاس‌ها را کاهش می‌دهد و از یک روش تنظیم جدید برای مقابله با منفرد شدن ماتریس پراکندگی درون-کلاسی استفاده می‌کند [۱۵]. اگر $x_l^{(i)}$ نمونه‌ی لام از کلاس λ باشد، دو نزدیکترین نسبت، برای آن در نظر گرفته می‌شود. $SNP(x_l^{(i)})$ که نزدیکترین نسبت از کلاس خودش و $ONP(x_l^{(i)})$ که نزدیکترین نسبت اختصاص یافته به آن از کلاس‌های دیگر است و به صورت زیر در نظر گرفته می‌شوند:

$$SNP(x_l^{(i)}) = \{x_{0NN}^{(i)}, x_{1NN}^{(i)}, x_{2NN}^{(i)}, \dots, x_{sNN}^{(i)}\} \quad (12)$$

$$ONP(x_l^{(i)}) = \{x_{1NN}^{(j)}, x_{2NN}^{(j)}, \dots, x_{oNN}^{(j)}\} \quad (13)$$

$$J(w) = \frac{w^T S_b^{\Phi} w}{w^T S_w^{\Phi} w} \quad (4)$$

که w یکی از بردارها در ماتریس تبدیل W است. S_b^{Φ} و S_w^{Φ} به ترتیب ماتریس پراکندگی بین-کلاسی و درون-کلاسی در فضای ویژگی \mathbb{H} هستند که به صورت زیر محاسبه می‌شوند:

$$S_b = \sum_{k=1}^c n_k (m_k^{\Phi} - m^{\Phi})(m_k^{\Phi} - m^{\Phi})^T \quad (5)$$

$$S_w = \sum_{k=1}^c \left(\sum_{i=1}^{n_k} (\Phi(x_{i,k}) - m_k^{\Phi})(\Phi(x_{i,k}) - m_k^{\Phi})^T \right) \quad (6)$$

که n_k تعداد نمونه‌های آموزشی کلاس k ام، m_k^{Φ} میانگین نمونه‌های آموزشی در کلاس k ام، m^{Φ} میانگین کل نمونه‌های آموزشی و $\Phi(x_{i,k})$ ، Φ ، آامین نمونه‌ی آموزشی در کلاس k ام است. Φ نشان دهنده‌ی این است که همه‌ی نمونه‌ها در فضای ویژگی \mathbb{H} هستند و محاسبات در این فضا انجام می‌شود. مشخص است که محاسبات نگاشت داده‌ها و سپس اجرای LDA از نظر محاسباتی پیچیده و زمان‌بر است. حتی ممکن است بعد فضای \mathbb{H} بی‌نهایت باشد. به همین دلیل، الگوریتم با جایگزینی ضرب‌های داخلی در فضای ویژگی جدید با تابع هسته مناسب $k(x_i, x_j) = \langle \varphi(x_i), \varphi(x_j) \rangle$ بازنویسی می‌شود. روش GDA هم مانند LDA تنها قادر به استخراج $c-1$ ویژگی می‌باشد ولی دارای کارایی بهتری نسبت به آن می‌باشد. از طرف دیگر برای داشتن کارکرد خوب، انتخاب نوع مناسب تابع کرنل در آن حائز اهمیت است.

استخراج ویژگی وزن دار غیر پارامتریک (NWFE)

این روش نظارت شده و غیر پارامتریک است و نسبت به LDA برتری دارد. ایده‌ی اساسی روش NWFE، اختصاص وزن‌های متفاوت به هر نمونه برای محاسبه‌ی میانگین وزن دار و تعریف ماتریس‌های پراکندگی درون-کلاسی و بین-کلاسی غیر پارامتریک جدید برای استخراج بیش از $c-1$ ویژگی است [۱۴]. ماتریس‌های پراکندگی در این روش به شکل زیر تعریف شده‌اند:

$$S_b = \sum_{i=1}^c p_i \sum_{j=1}^c \sum_{l=1}^{n_i} \frac{\lambda_l^{(i,j)}}{n_i} (x_l^{(i)} - M_j(x_l^{(i)})) (x_l^{(i)} - M_j(x_l^{(i)}))^T \quad (7)$$

$$\mathbf{X} \approx \mathbf{W}\mathbf{P}\mathbf{X} \quad (18)$$

تابع هدف مورد استفاده به این منظور، به شکل زیر تعریف می‌شود:

$$\arg \min_{\mathbf{W}, \mathbf{P} \geq 0} \|\mathbf{X} - \mathbf{W}\mathbf{P}\mathbf{X}\|_F^2 \quad (19)$$

که $\|\cdot\|_F$ نرم فروبنیوس است.

برای حفظ ساختار محلی داده‌ی ابرطیفی، فرض می‌کنیم که: اگر بتوان یک نقطه از داده (x_i) را بر حسب همسایه‌هایش در فضای ورودی بازنویسی کرد، همان نقطه را می‌توان با استفاده از همان ضرایب توسط همسایه‌هایش در زیرفضای با بعد کمتر بازسازی نمود. برای این منظور، یک ساختار گرافی برای داده در نظر می‌گیریم. اگر $N(x_i)$ مجموعه‌ی k نزدیک‌ترین همسایه‌های x_i باشند، هر کدام از همسایه‌های x_i را $x_j \in N(x_i)$ در نظر گرفته و ضریب بازسازی D_{ij} با تابع هدف زیر، به دست می‌آید:

$$\min \sum_i \|x_i - \sum_j D_{ij} x_j\|^2 \quad (20)$$

دو قید زیر، در رابطه‌ی فوق در نظر گرفته می‌شود:

۱- اگر هیچ لبه‌ای میان x_i و x_j در گراف موجود نباشد، $D_{ij} = 0$

۲- اگر x_i با ترکیب همسایه‌هایش، تقریب زده شود، $\sum_{j=1}^k D_{ij} = 1$

همان‌طور که در فضای اولیه، هر نمونه با استفاده از همسایه‌هایش قابل بازسازی است، به همان ترتیب در فضای کاهش یافته نیز، آن نمونه از همسایه‌هایش قابل بازسازی خواهد بود:

$$\begin{aligned} \emptyset &= \sum_i \left\| y_i - \sum_j D_{ij} y_j \right\|^2 = \|\mathbf{Y}(\mathbf{I} - \mathbf{D})\|_F^2 \\ &= \text{Tr}(\mathbf{Y}(\mathbf{I} - \mathbf{D})(\mathbf{I} - \mathbf{D})^T \mathbf{Y}^T) \\ &= \text{Tr}(\mathbf{P}\mathbf{X}\mathbf{M}\mathbf{X}^T \mathbf{P}^T) \end{aligned} \quad (21)$$

در رابطه‌ی فوق، y_i بیان بعد کمتر x_i در زیرفضای کاهش یافته، $y_i = \mathbf{P}x_i$ و \mathbf{I} ماتریس همانی است. همچنین $\mathbf{M} = (\mathbf{I} - \mathbf{D})(\mathbf{I} - \mathbf{D})^T$ می‌باشد.

میانگین محلی هر نسبت به شکل زیر تعریف می‌شود:

$$m_i(x_i^{(i)}) = \frac{1}{s+1} \sum_{t=0}^s x_{tNN}^{(i)} \quad (14)$$

$$m_j(x_i^{(i)}) = \frac{1}{o} \sum_{t=1}^o x_{tNN}^{(j)} \quad (15)$$

ماتریس پراکندگی بین-کلاسی در روش DNP به شکل زیر تعریف می‌شود:

$$\begin{aligned} \mathbf{S}_b &= \frac{1}{2} \sum_{l=1}^c p_l \sum_{\substack{j=1 \\ j \neq l}}^c p_j \sum_{i=1}^{n_i} \alpha_l^{(i,j)} (m_i(x_i^{(l)}) - m_j(x_i^{(l)})) \\ &\times (m_i(x_i^{(l)}) - m_j(x_i^{(l)}))^T \end{aligned} \quad (16)$$

که در رابطه‌ی بالا داریم:

$$\alpha_l^{(i,j)} = \frac{d(m_i(x_i^{(l)}), m_j(x_i^{(l)}))^{-1}}{\sum_{t=1}^{n_i} d(m_i(x_i^{(l)}), m_j(x_i^{(l)}))^{-1}} \quad (17)$$

در رابطه‌ی فوق، $\alpha_l^{(i,j)}$ تابع وزن برای تاکید ساختار مرزی توزیع است. نمونه‌ای که فاصله‌ی بین دو میانگین محلی آن کوچکتر است، در معیار جدایی‌پذیری بین کلاسی، وزن بیشتری به آن اختصاص داده می‌شود. در [۱۵]، DNP با NWFE در سایزهای مختلف نمونه‌ی آموزشی مقایسه شده است. ما تنها حالت تعداد نمونه‌ی آموزشی محدود، که مورد علاقه‌ی ما در این مقاله است، را بررسی می‌کنیم. برای داده‌ی ابرطیفی ایندیانا، بالاترین دقت به دست آمده در [۱۵] با ۲۰ نمونه‌ی آموزشی در هر کلاس و با استفاده از ماشین بردار پشتیبان (SVM^{15}) برای هر دو روش DNP و NWFE تقریباً یکسان و برابر ۰.۶۷ به دست آمده است.

عامل‌بندی ماتریس غیرمنفی تبدیل متعامد با حفظ همسایگی (NPOPNNMF)

برای به دست آوردن ماتریس تبدیل NPOPNNMF از دو عبارت در حل مسئله‌ی بهینه سازی آن استفاده می‌شود که یک عبارت مربوط به عامل بندی ماتریس غیر منفی تبدیل و عبارت دیگر مربوط به حفظ همسایگی محلی می‌باشد [۱۶]. با فرض ماتریس ورودی $\mathbf{X}_{d \times N}$ که ستون‌های آن را بردارهای داده‌ی d بعدی تشکیل داده‌اند، سعی در به دست آوردن ماتریس تبدیل $\mathbf{P}_{n \times d}$ و ماتریس غیرمنفی $\mathbf{W}_{d \times n}$ هستیم، به گونه ای که:

و درون-کلاسی (S_w) در روش پیشنهادی به شکل زیر تعریف می‌شوند:

$$S_b = \sum_{k=1}^c \sum_{i=1}^{n_k} n_k d(\mathbf{x}_{i,k}, \mathbf{m})(\mathbf{x}_{i,k} - \mathbf{m})(\mathbf{x}_{i,k} - \mathbf{m})^T \quad (23)$$

$$S_w = \sum_{k=1}^c \sum_{j=1}^{n_k} \sum_{i=1}^{n_k} d(\mathbf{x}_{i,k}, \mathbf{x}_{j,k})(\mathbf{x}_{i,k} - \mathbf{x}_{j,k})(\mathbf{x}_{i,k} - \mathbf{x}_{j,k})^T \quad (24)$$

n_k و m به ترتیب تعداد نمونه‌های آموزشی کلاس k ام و میانگین کل نمونه‌های آموزشی هستند. $\mathbf{x}_{i,k}$ نمونه‌ی i ام در کلاس k ام و c تعداد کلاس‌ها است. هم‌چنین $d(\mathbf{a}, \mathbf{b})$ به معنای فاصله‌ی اقلیدسی بین \mathbf{a} و \mathbf{b} می‌باشد. به دلیل محدود نبودن رتبه‌ی ماتریس پراکندگی بین-کلاسی، روش پیشنهادی قادر به استخراج هر تعداد دلخواه ویژگی مورد نیاز می‌باشد و تعداد ویژگی‌های استخراج شده توسط آن محدود به تعداد کلاس‌ها نیست. به علاوه، ما برای به دست آوردن ماتریس تبدیل استخراج ویژگی $W_{d \times n} = (w_1, w_2, \dots, w_n)$ که از بردار $\mathbf{x}_i \in \mathbb{R}^d$ در فضای اولیه با بعد بالا، ما را به $\mathbf{z}_i \in \mathbb{R}^n$ در فضای ثانویه با بعد کمتر می‌رساند ($\mathbf{z}_i = W^T \mathbf{x}_i$) به جای حداکثر کردن معیار $tr(S_w^{-1} S_b)$ از حداکثر نمودن معیار $tr(S_b - S_w)$ استفاده می‌کنیم. در نتیجه، به دلیل نیاز نداشتن به محاسبه‌ی معکوس ماتریس پراکندگی درون-کلاسی، با مشکل منفرد شدن S_w در نمونه‌های آموزشی محدود نیز، برخورد نخواهیم کرد.

توجه شود که روش استخراج ویژگی LDA برای محاسبه‌ی ماتریس‌های پراکندگی از میانگین کلی^{۱۶} برای کسب اطلاعات کلی کلاس‌ها استفاده می‌کند، در حالی که روش‌هایی مانند NWFE و DNP از میانگین‌های محلی^{۱۷} برای پیدا کردن اطلاعات مرزی کلاس‌ها استفاده می‌کنند. روش پیشنهادی بر خلاف روش‌های فوق، تنها برای تولید ماتریس پراکندگی بین-کلاسی نیاز به محاسبه‌ی میانگین کلی داشته و برای تولید ماتریس پراکندگی درون-کلاسی، نیاز به محاسبه‌ی هیچ‌گونه میانگین کلی یا محلی ندارد و از خود نمونه‌های آموزشی برای محاسبه‌ی ماتریس S_w بهره می‌برد. در نتیجه، در شرایطی که نمونه‌ی آموزشی محدودی در دسترس است و تخمین دقیقی از

روش پیشنهادی در [۱۶] از ترکیب دو رابطه‌ی بهینه‌سازی (۱۹) و (۲۱) به صورت زیر، به عنوان تابع هدف استفاده می‌کند و به منظور حفظ تعامد در ماتریس تبدیل \mathbf{P} ، قید $\mathbf{P}^T \mathbf{P} = \mathbf{I}$ را به آن اضافه می‌نماید:

$$F = \arg \min_{W, P \geq 0} \|X - WPX\|_F^2 + \lambda Tr(PXMX^T P^T)$$

$$\text{subject to } \mathbf{P}^T \mathbf{P} = \mathbf{I} \quad (22)$$

که $\lambda \geq 0$ یک پارامتر تنظیم است.

روش NPOPNNMF می‌تواند به شکل نظارت شده یا بدون نظارت پیاده‌سازی شود. در فرم نظارت شده، بین هر دو نمونه در ساختار گراف زمانی لبه‌ای قرار می‌گیرد که علاوه بر همسایگی، آن دو نمونه هم‌کلاس باشند و در فرم بدون نظارت، لبه بین نمونه‌های همسایه قرار می‌گیرد. مرجع [۱۶] دو روش NPOPNNMF و NWFE را برای داده‌ی ابرطیفی واشنگتن در ۲۰ نمونه‌ی آموزشی با استفاده از طبقه‌بند نزدیک‌ترین همسایه، مقایسه کرده است. بالاترین دقت متوسط به دست آمده در [۱۶] برای NWFE برابر ۰.۹۸ و ۰.۸۶٪ و برای روش NPOPNNMF در فرم بدون نظارت ۰.۹۸ و ۰.۵۷٪ و در فرم نظارت شده ۰.۹۹٪ به دست آمده است. همان طور که از نتایج حاصله دیده می‌شود، با استفاده از روش NPOPNNMF به شکل نظارت شده، تنها ۰.۰۴٪ بهبود نسبت به NWFE حاصل شده که قابل توجه نمی‌باشد. توجه کنید که روش‌های نظارت شده‌ی DNP و NPOPNNMF، در سایز کوچک نمونه‌ی آموزشی دارای نتایجی مشابه NWFE می‌باشند و لذا از بین سه روش NWFE، DNP و NPOPNNMF تنها روش پرکاربرد NWFE برای مقایسه با روش پیشنهادی در این مقاله انتخاب شده است. علاوه بر NWFE روش پیشنهادی با روش‌های معمول و پرکاربرد LDA و GDA نیز مقایسه شده است.

روش پیشنهادی

در این قسمت به توصیف روش پیشنهادی خود می‌پردازیم و در قسمت بعدی آن را با تعدادی از روش‌های پرکاربرد استخراج ویژگی مورد مقایسه قرار خواهیم داد. روش استخراج ویژگی پیشنهادی نظارت شده و غیرپارامتریک است و سعی در فشرده‌سازی نمونه‌های هم‌کلاس و جدانمودن نمونه‌های غیر هم‌کلاس می‌نماید. ماتریس‌های پراکندگی بین-کلاسی (S_b)

^{۱۶} global
^{۱۷} local

تصویر 512×217 بوده و حاوی ۱۶ کلاس است. داده‌ی ابرطیفی سوم، یک تصویر شهری مربوط به مرکز شهر پاولیا^{۲۰} در ایتالیا است. این تصاویر با استفاده از طیف‌سنج تصویربردار سیستم نوری انعکاسی (ROSI^{۲۱}) اخذ شده است. تصویر مرکز پاولیا حاوی 1096×715 بوده و حاوی ۱۰۲ باند طیفی و ۹ کلاس است.

طبقه‌بند مورد استفاده در انجام آزمایش‌ها، طبقه‌بند معروف و پرکاربرد SVM می‌باشد [۱۷] که در حالت چندکلاس با استفاده از روش یکی در برابر یکی و با تابع هسته‌ی شعاعی $k(x, y) = \exp(-\gamma \|x - y\|^2)$ ، RBF^{۲۲} LIBSVM [۱۸] پیاده‌سازی شده است. طبقه‌بند SVM با تابع هسته‌ی RBF دارای ۲ پارامتر است: پارامتر جریمه C و پهنای (انحراف معیار) تابع هسته‌ی RBF که با γ نشان داده شده است. برای هر داده‌ی ابرطیفی، با استفاده از یک ارزیابی متقاطع پنج مرتبه‌ای^{۲۳}، بهترین مقدار برای پارامتر C با یک جستجو از میان مجموعه $\{10^{-1}, 10^0, 10^1, 10^2, 10^3\}$ و بهترین مقدار برای پارامتر γ از میان مجموعه $\{10^{-3}, 10^{-2}, 10^{-1}, 10^0, 10^1\}$ انتخاب شده است. معیار ارزیابی طبقه‌بندی، دقت متوسط، دقت کل و اعتبار متوسط طبقه‌بندی است. دقت متوسط، میانگین دقت به دست آمده برای تمام کلاس‌ها است و دقت کل درصد کل پیکسل‌هایی است که درست طبقه‌بندی شده‌اند. اعتبار متوسط نیز برابر میانگین اعتبار به دست آمده برای تمام کلاس‌ها می‌باشد. اعتبار برای هر کلاس این‌گونه تعریف شده است: تعداد نمونه‌هایی که درست طبقه‌بندی شده‌اند، تقسیم بر تعداد کل نمونه‌هایی که به آن کلاس تعلق گرفته‌اند. در تمام آزمایش‌ها، به دلیل بررسی کارایی روش پیشنهادی در تعداد نمونه‌های آموزشی محدود، تنها ۱۶ نمونه‌ی برجسته‌ی خورده از هر کلاس به عنوان نمونه‌ی آموزشی و بقیه به عنوان نمونه‌ی آزمایشی در نظر گرفته شده‌اند. نمونه‌های آموزشی به صورت تصادفی از کل صحنه انتخاب گردیده‌اند. هر آزمایش ۱۰ مرتبه تکرار شده است و متوسط نتایج به دست آمده در مقاله گزارش شده‌اند. هم‌چنین به دلیل مقایسه‌ی منصفانه در هر آزمایش، برای

میانگین کلی یا محلی قابل محاسبه نخواهد بود، استفاده از روش استخراج ویژگی پیشنهادی، کارا تر از سایر روش‌های موجود خواهد بود.

علاوه بر مزایای فوق، روش پیشنهادی از نظر محاسباتی نیز کارا بوده و نسبت به روش غیرپارامتریک NWF^{۲۴} به زمان محاسباتی بسیار کمتری نیاز دارد. با نگاهی به روابط (۷) تا (۱۰) به سادگی می‌توان مشاهده کرد که برای محاسبه‌ی ماتریس‌های S_w و S_b در NWF^{۲۴} نیاز به انجام محاسبات زیاد و حلقه‌های تو در تو می‌باشد. در حالی که روابط (۲۳) و (۲۴) مربوط به روش پیشنهادی، بسیار ساده می‌باشند و جز محاسبه‌ی کم هزینه‌ی میانگین کلی در ماتریس S_b و وزن ماتریس‌ها که فاصله‌ی اقلیدسی بین نمونه‌ها یا بین نمونه و میانگین کلی می‌باشد، نیاز به محاسبات پیچیده‌تری ندارد.

روش پیشنهادی نسبت به روش‌های استخراج ویژگی بیان شده، در شرایطی که نمونه‌های آموزشی محدود در دسترس است، از نظر دقت طبقه‌بندی برتری دارد. در قسمت بعدی، با پیاده‌سازی روش پیشنهادی و تعدادی از روش‌های استخراج ویژگی بر روی سه مجموعه داده‌ی ابرطیفی، به بررسی و مقایسه‌ی آن‌ها از نظر سرعت و دقت می‌پردازیم.

نتایج آزمایش‌ها

سه مجموعه داده از تصاویر ابرطیفی برای انجام آزمایش‌ها در این بخش مورد استفاده قرار گرفته‌اند. داده‌ی ابرطیفی اول، توسط سنجنده‌ی AVIRIS از یک سایت آزمایشی در شمال شرقی ایالت ایندیانا^{۱۸} گرفته شده است. این تصویر دارای بعد مکانی 145×145 بوده و حاوی ۱۶ کلاس مختلف از منطقه‌ی کشاورزی- جنگلی می‌باشد. ۱۰ کلاس از این داده برای انجام آزمایش‌ها مورد استفاده قرار گرفته‌اند. تصویر ابرطیفی ایندیانا دارای ۲۲۰ باند طیفی است که ۲۰ باند جذب آب آن حذف شده و ۲۰۰ باند باقی‌مانده در آزمایش‌ها مورد استفاده قرار گرفته‌اند. دومین داده‌ی ابرطیفی، نیز توسط سنجنده‌ی AVIRIS از دره‌ی سالیناس^{۱۹} در کالیفرنیا تصویر برداری شده است و حاوی ۲۲۴ باند است که پس از حذف ۲۰ باند جذب آب، ۲۰۴ باند طیفی باقی مانده در آزمایش‌ها استفاده شده‌اند. این

^{۲۰} Pavia center
^{۲۱} Reflective Optics System Imaging Spectrometer
^{۲۲} Radial basis function
^{۲۳} fivefold cross validation

^{۱۸} Indian
^{۱۹} Salinas

آزمایش‌ها می‌توان مشاهده کرد که با استفاده از سایز کوچک نمونه‌ی آموزشی، روش پیشنهادی دارای دقت و سرعت بهتری نسبت به سایر روش‌های بیان شده‌ی استخراج ویژگی می‌باشد. روش LDA نسبت به همه‌ی روش‌ها سریع‌تر است، ولی کمترین دقت طبقه‌بندی را فراهم می‌آورد. بعد از روش پیشنهادی، NWFE دارای دقت طبقه‌بندی مناسبی است، ولی نسبت به سایر روش‌ها، به زمان محاسباتی بیشتری نیاز دارد.

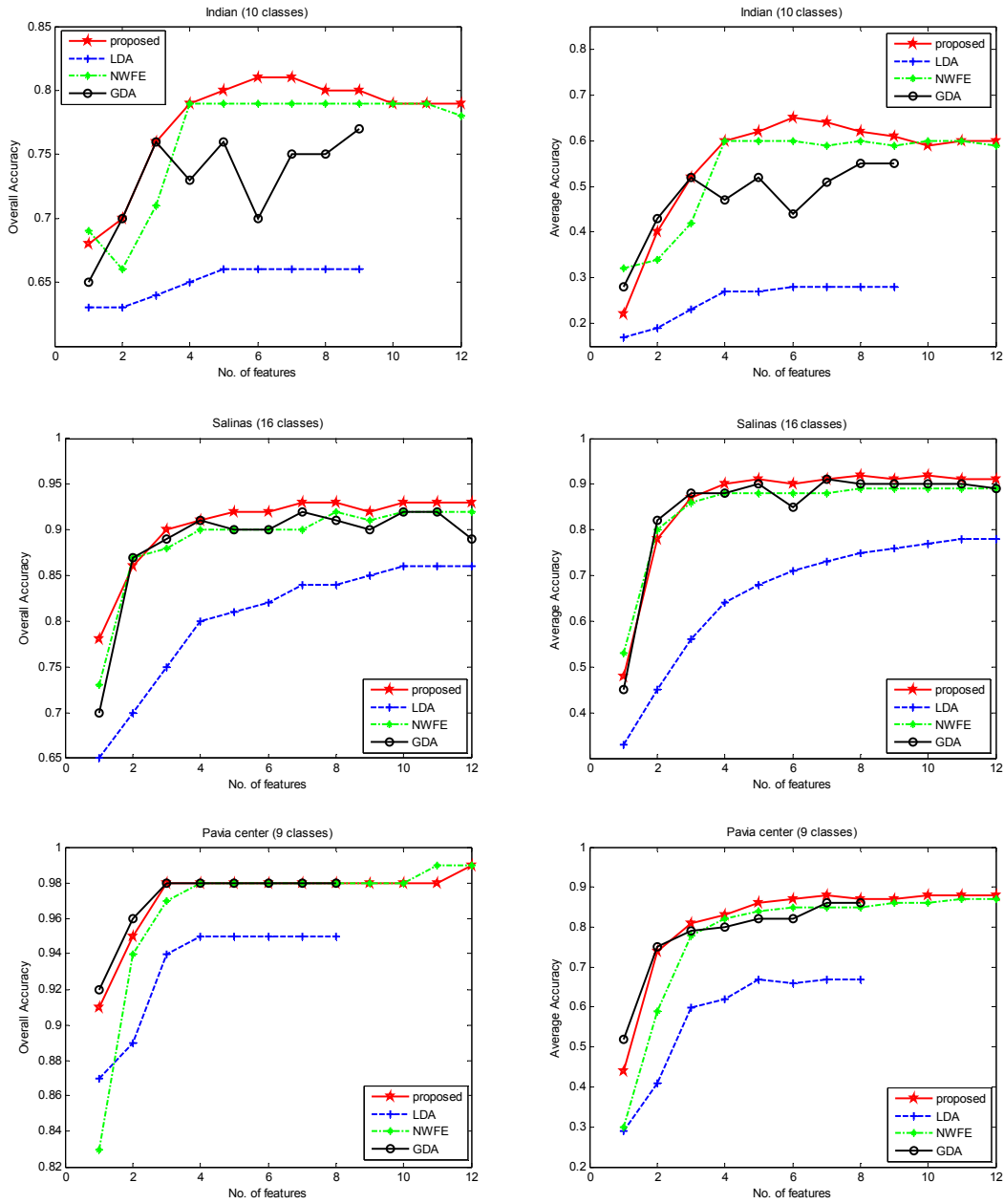
تصویر ایندیانان به دلیل شرایط خاص منطقه مورد بررسی با مشکلات زیادی برای طبقه‌بندی مواجه است. مهمترین مشکل این است که گیاهان موجود در این منطقه، که بیشتر ذرت و سویا هستند، در اوایل دوره رشد خود بوده‌اند. از طرفی چون منطقه مورد بررسی یک سایت آزمایشی می‌باشد، مقداری از محصولات کاشته شده در سال‌های قبل، به صورت خاشاک بر روی سطح زمین باقیمانده‌اند. نواحی منطقه مورد بررسی براساس مقدار خاشاک که از گیاهان سال‌های قبل بر روی زمین باقیمانده است، به سه دسته تقسیم می‌شود: ناحیه بدون شخم، ناحیه کم شخم و ناحیه با شخم کامل. به دلیل متفاوت بودن مواد موجود در سطح نواحی که گیاهان یکسان دارند، طیف‌های اندازه‌گیری شده برای یک گیاه در نواحی مختلف تصویر تفاوت بسیار زیادی با یکدیگر خواهند داشت. از طرفی بیشتر منطقه را زمین لخت و یا خاشاک گیاهان سال‌های قبل پوشانده است. این عوامل باعث می‌شوند که توزیع احتمال کلاس‌ها چند مده و با همپوشانی زیاد باشند. بنابراین میزان جدایی‌پذیری کلاس‌ها کم و طبقه‌بندی نظارت شده پیکسل‌ها در این تصویر کار بسیار دشواری خواهد بود. به دلایل ذکر شده، تفاوت کارایی روش‌های استخراج ویژگی با استفاده از نمونه‌های آموزشی محدود، برای این داده‌ی ابرطیفی مشخص‌تر از سایر داده‌ها می‌باشد. برای داده‌ی ایندیانان، دقت متوسط به دست آمده برای روش پیشنهادی نسبت به روش‌های LDA، NWFE و GDA به ترتیب ۰/۳۷، ۰/۵ و ۰/۲۱ بهبود یافته است. در این داده، LDA نسبت به روش پیشنهادی ۰/۲ برابر سریع‌تر و روش پیشنهادی نسبت به NWFE و GDA به ترتیب ۰/۳۶، ۰/۴۱ برابر و ۰/۱۲ برابر سریع‌تر می‌باشد.

همه‌ی روش‌های استخراج ویژگی، نمونه‌های آموزشی یکسانی مورد استفاده قرار گرفته‌اند.

نمودار به دست آمده‌ی دقت کل و دقت متوسط در برابر تعداد ویژگی‌های استخراج شده برای ۳ داده‌ی ابرطیفی ایندیانان، سالیانس و مرکز پاپوا در شکل ۱ نشان داده شده‌اند. با توجه به این نمودارها، می‌توان نکات زیر را استنباط کرد:

- ۱- روش استخراج ویژگی پیشنهادی نسبت به سایر روش‌ها، به طور کلی دارای کارایی بهتری می‌باشد.
- ۲- برای مجموعه داده‌ی ایندیانان، پس از روش استخراج ویژگی پیشنهادی که غیر پارامتریک است، روش غیر پارامتریک وزن‌دار NWFE دارای بهترین کارایی است و برای دو داده‌ی ابرطیفی دیگر، GDA و NWFE دارای کارایی نزدیک به هم می‌باشند.
- ۳- روش‌های LDA و GDA به دلیل محدودیت رتبه‌ی ماتریس پراکندگی بین-کلاسی، تنها قادر به استخراج 1- c ویژگی هستند که این تعداد ویژگی برای به دست آوردن دقت بهینه طبقه‌بندی ممکن است کافی نباشند.
- ۴- LDA نسبت به سایر روش‌های نظارت شده‌ی استخراج ویژگی به دلیل مشکل منفرد شدن ماتریس پراکندگی درون-کلاسی با نمونه‌های آموزشی محدود، دارای بدترین کارایی است.

بیشترین دقت طبقه‌بندی توسط روش استخراج ویژگی پیشنهادی برای داده‌ی ایندیانان به ازای ۶ ویژگی، برای داده‌ی سالیانس به ازای ۸ ویژگی و برای داده‌ی مرکز پاپوا به ازای ۷ ویژگی حاصل شد که دقت و اعتبار کلاس‌ها در این سه مورد، در جدول‌های ۱، ۲ و ۳ آورده شده‌اند. همین‌طور، نقشه‌های کلاس به دست آمده در این موارد به همراه نقشه کلاس واقعی (GTM^{۲۴}) سه تصویر در شکل‌های ۲، ۳ و ۴ دیده می‌شوند. علاوه بر دقت طبقه‌بندی، روش استخراج ویژگی پیشنهادی، از نظر سرعت اجرا نیز با سایر روش‌ها مقایسه گردیده است. برای بررسی و مقایسه‌ی زمان محاسباتی روش‌ها، برای هر داده‌ی ابرطیفی، فرآیند استخراج ویژگی ۱۰ مرتبه تکرار شده و متوسط زمان به دست آمده از آزمایش‌ها (برحسب ثانیه) در جدول ۴ گزارش شده است. با بررسی نتایج به دست آمده از



شکل ۱. دقت کل و دقت متوسط در برابر تعداد ویژگی‌های استخراج شده برای ۳ داده‌ی ابرطیفی ایندیانا، سالیناس و مرکز پاولیا

جدول ۱. دقت و اعتبار طبقه‌بندی کلاس‌ها در ۴ روش استخراج ویژگی برای داده‌ی ایندیانا با استفاده از ۶ ویژگی

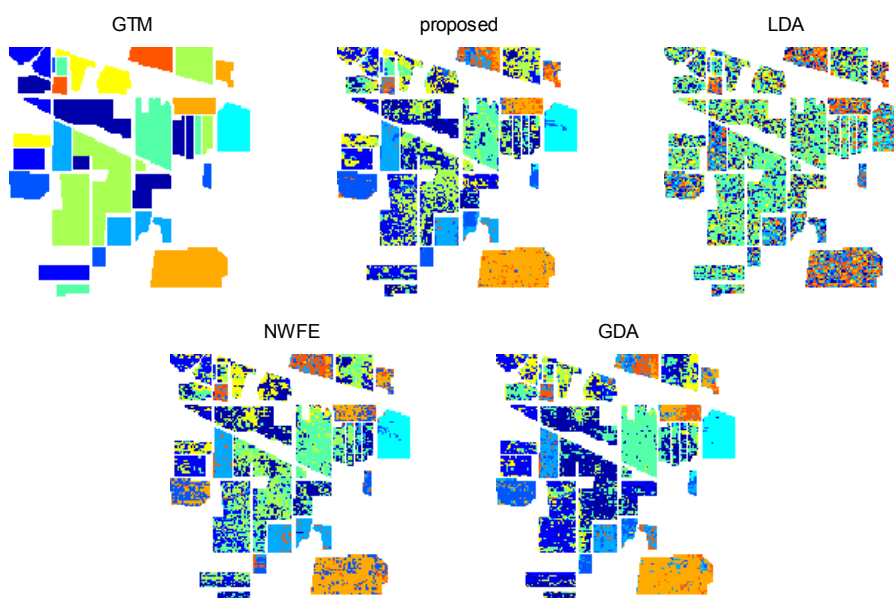
کلاس			Proposed		LDA		NWFE		GDA	
شماره	نام کلاس	#نمونه‌ها	دقت	اعتبار	دقت	اعتبار	دقت	اعتبار	دقت	اعتبار
1	ذرت-بدون شخم	1434	0.61	0.41	0.22	0.22	0.54	0.37	0.53	0.22
2	ذرت-کم شخم	834	0.55	0.50	0.17	0.16	0.43	0.48	0.14	0.11
3	سبزه-درختان	497	0.74	0.66	0.55	0.17	0.75	0.60	0.56	0.34
4	سبزه-چمنزار	747	0.81	0.84	0.16	0.29	0.90	0.75	0.40	0.76
5	کاه و خاشاک	489	0.98	0.95	0.53	0.71	0.99	1.00	0.98	0.93
6	سویا-بدون شخم	968	0.61	0.52	0.11	0.15	0.35	0.35	0.25	0.30
7	سویا-کم شخم	2468	0.44	0.72	0.26	0.34	0.39	0.68	0.02	0.76
8	سویا-شخم کامل	614	0.46	0.44	0.20	0.14	0.49	0.31	0.31	0.23
9	بیشه	1294	0.65	0.94	0.35	0.54	0.61	0.90	0.65	0.86
10	ساختمان-سبزه	380	0.60	0.32	0.22	0.11	0.55	0.33	0.55	0.28
دقت متوسط و اعتبار متوسط			0.65	0.63	0.28	0.28	0.60	0.58	0.44	0.48
دقت کل			0.81		0.66		0.79		0.70	

جدول ۲. دقت و اعتبار طبقه‌بندی کلاس‌ها در ۴ روش استخراج ویژگی برای داده‌ی سالییناس با استفاده از ۸ ویژگی

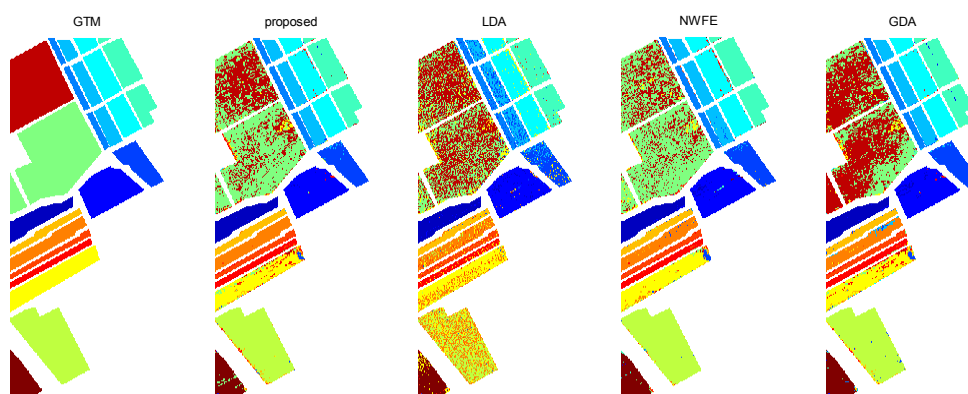
کلاس			Proposed		LDA		NWFE		GDA	
شماره	نام کلاس	#نمونه‌ها	دقت	اعتبار	دقت	اعتبار	دقت	اعتبار	دقت	اعتبار
1	بروکلی-سبزه-علف هرز ۱	2009	0.95	1.00	0.91	0.97	0.92	1.00	0.96	1.00
2	بروکلی-سبزه-علف هرز ۲	3726	0.92	0.97	0.95	0.95	0.89	0.95	0.95	0.98
3	زمین شخم شده	1976	0.89	0.94	0.68	0.59	0.88	0.93	0.83	0.90
4	زمین شخم شده-زیر-گاواهن	1394	0.98	0.98	0.85	0.81	0.94	0.97	0.99	0.92
5	زمین شخم شده-هموار	2678	0.97	0.98	0.53	0.79	0.97	0.93	0.93	0.98
6	کاه بن	3959	0.97	1.00	0.90	1.00	0.98	1.00	0.98	0.99
7	کرفس	3579	0.99	0.92	0.98	0.99	0.99	0.98	0.99	0.99
8	انگور	11271	0.69	0.79	0.52	0.67	0.62	0.72	0.74	0.65
9	خاک-تاکستان	6203	0.97	0.98	0.67	0.91	0.99	0.96	0.96	0.98
10	ذرت-سبزه پیر-علف هرز	3278	0.88	0.79	0.66	0.61	0.85	0.75	0.92	0.73
11	کاهو-۴ هفته	1068	0.93	0.77	0.81	0.97	0.86	0.94	0.93	0.78
12	کاهو-۵ هفته	1927	1.00	0.97	0.72	0.17	0.99	0.90	0.95	0.97
13	کاهو-۶ هفته	916	0.99	0.62	0.75	0.68	0.99	0.69	0.99	0.92
14	کاهو-۷ هفته	1070	0.89	0.92	0.75	0.84	0.88	0.96	0.91	0.77
15	تاکستان	7268	0.68	0.63	0.40	0.45	0.58	0.53	0.35	0.51
16	تاکستان-شاغولی-داربست	1807	0.98	0.97	0.89	0.99	0.93	0.79	0.95	0.98
دقت متوسط و اعتبار متوسط			0.92	0.89	0.75	0.77	0.89	0.88	0.90	0.88
دقت کل			0.93		0.84		0.92		0.91	

جدول ۳. دقت و اعتبار طبقه‌بندی کلاس‌ها در ۴ روش استخراج ویژگی برای داده‌ی مرکز پاپویا با استفاده از ۷ ویژگی

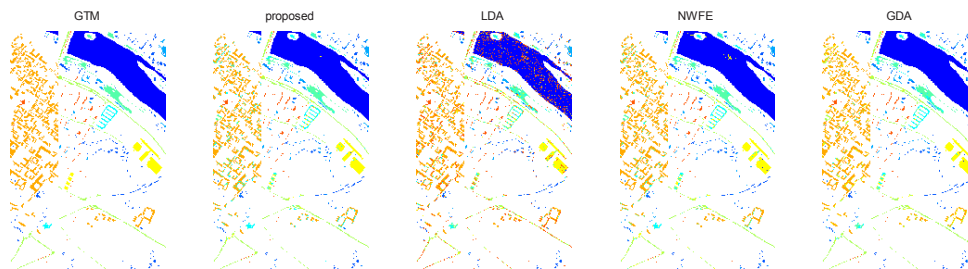
کلاس			Proposed		LDA		NWFE		GDA	
شماره	نام کلاس	#نمونه‌ها	دقت	اعتبار	دقت	اعتبار	دقت	اعتبار	دقت	اعتبار
1	آب	65971	1.00	1.00	0.85	1.00	1.00	1.00	1.00	1.00
2	درخت	7598	0.66	0.96	0.62	0.89	0.70	0.90	0.72	0.95
3	چمنزار	3090	0.92	0.52	0.74	0.71	0.80	0.52	0.90	0.55
4	آجر	2685	0.79	0.28	0.41	0.22	0.77	0.39	0.68	0.43
5	خاک بایر	6584	0.84	0.76	0.53	0.64	0.81	0.59	0.76	0.86
6	آسفالت	9248	0.91	0.90	0.51	0.66	0.78	0.96	0.91	0.75
7	قیر	7287	0.90	0.86	0.73	0.52	0.89	0.74	0.90	0.64
8	کاشی	42826	0.86	1.00	0.71	1.00	0.89	1.00	0.86	1.00
9	سایه	2863	1.00	1.00	0.95	0.11	1.00	0.96	1.00	1.00
دقت متوسط و اعتبار متوسط			0.88	0.81	0.67	0.64	0.85	0.79	0.86	0.80
دقت کل			0.98		0.95		0.98		0.98	



شکل ۲. نقشه‌های کلاس به دست آمده برای چهار روش استخراج ویژگی به همراه GTM داده‌ی ایندیانای برای ۶ ویژگی استخراج شده



شکل ۳. نقشه‌ی کلاس به دست آمده برای چهار روش استخراج ویژگی به همراه GTM داده‌ی سالیناس برای ۸ ویژگی استخراج شده



شکل ۴. نقشه‌ی کلاس به دست آمده برای چهار روش استخراج ویژگی به همراه GTM داده‌ی مرکز پاویا برای ۷ ویژگی استخراج شده

جدول ۴. زمان محاسباتی روش‌های استخراج ویژگی برای هر سه داده‌ی ابرطیفی (برحسب ثانیه)

	Proposed	LDA	NWFE	GDA
ایندیانا	2.50	0.85	91.02	31.40
سالیانس	5.59	5.48	211.49	59.83
مرکز پاویا	29.99	12.95	262.07	163

- Spectrometer (AVIRIS),” Remote Sens. Environ. , 1998, vol. 65, no. 3, pp. 227–248.
- [2] C. Dhanjal, S. R. Gunn, and J. Shawe-Taylor, “Efficient Sparse Kernel Feature Extraction Based on Partial Least Squares,” IEEE Trans. Pattern Analysis and Machine Intelligence, 2009, vol. 31, no. 8, pp. 1347–1361.
- [3] C. -I. Chang and S. Wang, “Constrained band selection for hyperspectral imagery,” IEEE Trans. Geosci. Remote Sens. , 2006 vol. 44, no. 6, pp. 1575–1585.
- [4] L. Zhang, Y. Zhong, B. Huang, J. Gong, and P. Li, “Dimensionality reduction based on clonal selection for hyperspectral imagery,” IEEE Trans. Geosci. Remote Sens. , 2007, vol. 45, no. 12, pp. 4172–4186.
- [5] L. Bruzzone, and C. Persello, “A Novel Approach to the Selection of Spatially Invariant Features for the Classification of Hyperspectral Images With Improved Generalization Capability,” IEEE Trans. Geosci. Remote Sens. , 2007, vol. 47, no. 9, pp. 3180–3191.
- [6] K. Fukunaga, Introduction to Statistical Pattern Recognition, 2nd ed, Academic Press, New York, 1990, pp. 400-506.
- [7] G. Baudat and F. Anouar, “Generalized discriminant analysis using a kernel approach,” Neural Comput. , vol. 12, no. 10, pp. 2385–2404, Oct. 2000.
- [8] X. Jia, B. C. Kuo, and M. Crawford, "Feature Mining for Hyperspectral Image

نتیجه‌گیری

در این مقاله یک روش نظارت شده‌ی غیر پارامتریک استخراج ویژگی برای طبقه‌بندی تصاویر ابرطیفی بیان شد که علاوه بر سادگی و زمان کم محاسباتی نسبت به سایر روش‌های غیر پارامتریک از قبیل NWFE دارای کارایی بهتری از نظر دقت طبقه‌بندی می‌باشد. مشکلاتی از قبیل منفرد شدن ماتریس پراکنندگی درون-کلاسی و محدودیت تعداد ویژگی‌های استخراجی که در روش معمول LDA وجود دارد، در روش پیشنهادی نیست و به این دلیل روش پیشنهادی، می‌تواند یک روش مناسب برای استخراج ویژگی با استفاده از نمونه‌های آموزشی محدود باشد. نتایج آزمایش‌ها بر روی سه داده‌ی ابرطیفی معمول و پرکاربرد، کارایی مناسب روش پیشنهادی را در مقایسه با تعدادی از روش‌های استخراج ویژگی نظارت شده‌ی معمول نشان می‌دهد.

مراجع

- [1] R. O. Green, M. L. Eastwood, C. M. Sarture, T. G. Chrien, M. Aronsson, B. J. Chippendale, J. A. Faust, B. E. Pavri, C. J. Chovit, M. S. Solis, M. R. Olah, and O. Williams, “Imaging spectroscopy and the Airborne Visible/Infrared Imaging

- [14] B. C. Kuo and D. A. Landgrebe, "Nonparametric weighted feature extraction for classification," *IEEE Trans. Geosci. Remote Sens.*, 2004, vol. 42, no. 5, pp. 1096-1105.
- [15] H. -Y. Huang, and B. -C. Kuo, "Double Nearest Proportion Feature Extraction for Hyperspectral-Image Classification," *IEEE Trans. Geosci. Remote Sens.*, 2010, vol. 48, no. 11, pp. 4034-4046.
- [16] J. Wen, Z. Tian, X. Liu, and Wei Lin, "Neighborhood Preserving Orthogonal PNM Feature Extraction for Hyperspectral Image Classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2013, vol. 6, no. 2, pp. 759-768.
- [17] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, 4th Ed., Elsevier Academic Press, 2009, San Diego, California, pp. 119-127.
- [18] C. Chang and C. Lin, LIBSVM—A Library for Support Vector Machines, 2008. [Online]. Available: <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- Classification," *Proceedings of the IEEE*, 2013, vol. 101, no. 3, pp. 676-697.
- [9] S. Kumar, J. Ghosh, and M. M. Crawford, "Best-Bases Feature Extraction Algorithms for Classification of Hyperspectral Data," *IEEE Trans. Geosci. Remote Sens.*, 2001, vol. 39, no. 7, pp. 1368-1379.
- [10] W. Liao, A. Pižurica, P. Scheunders, W. Philips, and Y. Pi, "Semisupervised Local Discriminant Analysis for Feature Extraction in Hyperspectral Images," *IEEE Trans. Geosci. Remote Sens.*, 2013, vol. 51, no. 1, pp. 184-198.
- [11] C. Cariou, K. Chehdi, and S. Le Moan, "BandClust: An Unsupervised Band Reduction Method for Hyperspectral Remote Sensing," *IEEE Geosci. Remote Sens. Lett.*, 2011, vol. 8, no. 3, pp. 565-569.
- [12] M. Kamandar and H. Ghassemian, "Linear Feature Extraction for Hyperspectral Images Based on Information Theoretic Learning," *IEEE Geoscience and Remote Sensing Lett.*, 2013, vol. 10, no. 4, pp. 702 - 706.
- [13] M. J. Mendenhall, and E. Merényi, "Relevance-Based Feature Extraction for Hyperspectral Images," *IEEE Trans. Neural Network*, 2008, vol. 19, no. 4, pp. 658-672.