

طبقه‌بندی تصاویر ابرطیفی با استفاده از تحلیل ممیز خطی فضای خوشه و مجموعه نمونه‌های آموزشی کوچک

مریم ایمانی و حسن قاسمیان

می‌شود. فرایند انتخاب ویژگی به معنای انتخاب زیرمجموعه مناسبی از ویژگی‌های اولیه و دورریختن مابقی ویژگی‌هاست. یک فایده انتخاب ویژگی، حفظ معنای فیزیکی ویژگی‌های انتخاب شده است. هدف فرایند انتخاب ویژگی، انتخاب مؤثرترین زیرمجموعه از ویژگی‌ها بر طبق یک تابع معیار و یک الگوریتم جستجو است [۸] و [۹]. انتخاب ویژگی با استخراج ویژگی که به معنای تبدیل و تصویر داده از فضای اولیه به فضای ثانویه دیگری (معمولاً با بعد کمتر) می‌باشد، متفاوت است. هدف اصلی ما در این مقاله، استخراج ویژگی می‌باشد.

استخراج ویژگی می‌تواند با استفاده از نمونه‌های آموزشی و به صورت نظارت‌شده انجام شود، مانند روش معمول و کلاسیک تحلیل ممیز خطی (LDA) [۱۰] یا می‌تواند به صورت بدون نظارت (بدون استفاده از نمونه‌های آموزشی) انجام گیرد. معروف‌ترین روش بدون نظارت استخراج ویژگی، تحلیل مؤلفه‌های اصلی (PCA) [۱۰] می‌باشد. دسته اول اشاره شده (روش‌های بانظارت) به دلیل در نظر گرفتن جدایی‌پذیری کلاس‌ها مناسب برای کاربردهای طبقه‌بندی هستند اما دسته دوم (روش‌های بدون نظارت) تمییز 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. علاوه بر این دسته‌بندی می‌توان یک گروه‌بندی دیگر برای روش‌های استخراج ویژگی در نظر گرفت: روش‌های خطی و غیر خطی. روش‌های غیر خطی معمولاً از تابع هسته (کرنل) برای بسط مدل خطی به مدل غیر خطی استفاده می‌کنند. با استفاده از یک تابع هسته مناسب برای تبدیل داده، جدایی‌پذیری کلاس‌ها در فضای کرنل می‌تواند افزایش یابد. از جمله این روش‌ها می‌توان به تحلیل ممیز تعمیم‌یافته (GDA) [۱۱] اشاره کرد که بسط غیر خطی روش LDA است. روش‌هایی همچون LDA به علت منفرد شدن ماتریس پراکندگی درون کلاسی، دارای کارایی ضعیفی با استفاده از تعداد نمونه‌های آموزشی محدود هستند. GDA نسبت به LDA دارای کارایی بهتری است. استخراج ویژگی وزن‌دار غیر پارامتریک (NWF) برای حل مشکلات LDA در [۱۲] معرفی شده است. NWF با محاسبه میانگین وزن‌دار، ماتریس‌های پراکندگی پارامتریک و غیر پارامتریک جدیدی تعریف می‌کند. جاسازی حفظ همسایگی (NPE) و تصویر حفظ محلی (LPP) دو روش معروف و مؤثر از مجموعه روش‌های یادگیری خمینه^۱ هستند که برای حفظ ساختار محلی داده استفاده می‌شوند. هر کدام از این دو روش می‌توانند به شکل بدون نظارت و یا با نظارت پیاده‌سازی شوند. روش LPP ساختار محلی داده را با استفاده از یک گراف مجاورت حفظ می‌کند [۱۳] در حالی که NPE خطای بازسازی را حداقل می‌کند [۱۴]. لازم به ذکر است که روش‌های NPE و LPP که هدف اصلی آنها حفظ

چکیده: امروزه، تصاویر ابرطیفی به دست آمده از سنجنده‌های از راه دور، امکان تمییز بین کلاس‌ها با جزئیات بیشتر را فراهم آورده‌اند. بعد بالای داده ابرطیفی از یک سو و کمبود تعداد نمونه‌های آموزشی از سوی دیگر، سبب ایجاد مشکلاتی در طبقه‌بندی تصاویر ابرطیفی می‌شود. از آنجایی که جمع‌آوری داده‌های آموزشی سخت و زمان‌بر است، کاهش تعداد باندهای طیفی به بهبود کارایی طبقه‌بند کمک شایانی خواهد نمود. بنابراین استخراج ویژگی تصاویر ابرطیفی، یکی از مراحل مهم پیش‌پردازش این نوع داده‌ها محسوب می‌شود. در شرایطی که تعداد نمونه آموزشی در دسترس محدود است، روش‌های استخراج ویژگی معمول همچون LDA دارای کارایی خوبی نخواهند بود. در این مقاله، یک روش استخراج ویژگی نظارت‌شده معرفی گردیده که بدون نیاز به تعیین برچسب نمونه‌های آزمایشی، با انجام خوشه‌بندی بر روی آنها و یافتن رابطه بین نمونه‌های آموزشی و آزمایشی، از قدرت نمونه‌های بدون برچسب به همراه نمونه‌های آموزشی برای تولید ماتریس‌های پراکندگی درون کلاسی و بین کلاسی استفاده می‌کند. این کار دقت طبقه‌بندی را خصوصاً برای داده‌های ابرطیفی چندمنده افزایش خواهد داد. روش پیشنهادی که تحلیل ممیز خطی فضای خوشه (CSLDA) نامیده می‌شود، با روش‌های پرکاربرد استخراج ویژگی نظارت‌شده همچون LDA، GDA و NWF مقایسه شده است. نتایج به دست آمده بر روی تصاویر ابرطیفی واقعی از مناطق شهری و کشاورزی نشان می‌دهند که روش پیشنهادی با استفاده از تعداد محدود نمونه‌های آموزشی، دارای کارایی بهتری نسبت به سایر روش‌ها می‌باشد.

کلیدواژه: ابرطیفی، استخراج ویژگی، خوشه‌بندی، طبقه‌بندی، نمونه آموزشی.

۱- مقدمه

استفاده از تصاویر ابرطیفی به دلیل فراهم آوردن حجم عظیم اطلاعات طیفی، امکان طبقه‌بندی با دقت بالا و تمییز بین کلاس‌ها با جزئیات بیشتر را به خوبی فراهم می‌کند. افزایش تعداد زیاد باندهای طیفی از این جهت مفید و بارز است، ولی از طرف دیگر به دلیل محدودیت تعداد نمونه‌های آموزشی با مشکل پدیده هیوز مواجه خواهیم شد [۱]. راه حل‌های مختلفی برای حل مشکل تعداد نمونه‌های آموزشی محدود وجود دارد که از آن جمله می‌توان به این موارد اشاره کرد: استفاده از طبقه‌بند مناسب غیر پارامتریک که حساسیت کمتری به تعداد نمونه‌های آموزشی دارند [۲] و [۳]، استفاده از نمونه‌های آموزشی نیمه‌برچسب‌دار [۴] و کاهش تعداد ویژگی‌ها [۵] تا [۷]. کاهش ویژگی خود به دو گروه کلی انتخاب ویژگی و استخراج ویژگی تقسیم‌بندی

این مقاله در تاریخ ۲۵ فروردین ماه ۱۳۹۳ دریافت و در تاریخ ۵ اسفند ماه ۱۳۹۳ بازنگری شد.

مریم ایمانی، دانشکده مهندسی برق و کامپیوتر، دانشگاه تربیت مدرس، تهران، (email: maryam.imani@modares.ac.ir)

حسن قاسمیان، دانشکده مهندسی برق و کامپیوتر، دانشگاه تربیت مدرس، تهران، (email: ghassemi@modares.ac.ir)

استفاده می‌کند. اساس این روش بر مبنای بیشینه‌کردن نسبت ماتریس کواریانس بین کلاسی به ماتریس کواریانس درون کلاسی است. فرض می‌شود که $\{x_i\}_{i=1}^N$ و $\{z_i\}_{i=1}^N$ به ترتیب داده بعد بالا (داده اولیه) و داده بعد پایین (داده تبدیل شده) بوده و d و n به ترتیب، بعد داده اولیه و بعد فضای تبدیل یافته باشند. نگاشت داده از $x_i \in \mathbb{R}^d$ به $z_i \in \mathbb{R}^n$ با یک تبدیل خطی $z_i = W^T x_i$ انجام می‌شود. بردار تشکیل دهنده ماتریس تبدیل $W_{d \times n} = (w_1, w_2, \dots, w_n)$ که با w نمایش داده شده با حل مسئله بهینه‌سازی زیر به دست می‌آید

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

ماتریس‌های پراکنده‌گی درون کلاسی S_w و بین کلاسی S_b در LDA به شکل زیر تعریف می‌شوند

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

$$S_w = \sum_{k=1}^c \sum_{i=1}^{n_k} (x_{i,k} - m_k)(x_{i,k} - m_k)^T \quad (3)$$

n_k ، m_k ، m و $x_{i,k}$ به ترتیب تعداد نمونه‌های آموزشی کلاس k ام، میانگین کلاس k ام، میانگین کل نمونه‌های آموزشی و نمونه i ام در کلاس k ام بوده و 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. روش LDA یک روش معروف و پرکاربرد نظارت شده است ولی در عین حال دارای ایراداتی می‌باشد. از آن جمله به موارد زیر می‌توان اشاره کرد: با وجود نمونه‌های آموزشی محدود، ماتریس پراکنده‌گی درون کلاسی، منفرد شده و در این حالت، LDA کارایی ضعیفی خواهد داشت. به علاوه LDA زمانی خوب کار می‌کند که کلاس‌ها دارای توزیع نرمال باشند و اگر توزیع کلاس‌ها غیر نرمال یا چندمنه باشد، کارایی روش LDA چندان خوب نخواهد بود.

تحلیل ممیز تعمیم یافته (GDA): می‌توان روش استخراج ویژگی خطی LDA را از طریق نگاشت داده از فضای ورودی به فضای ویژگی با بعد بالاتر به روش غیر خطی GDA بسط داد. با استفاده از یک نگاشت Φ می‌توان نمونه $x \in \mathbb{R}^d$ را به فضای بعد بالاتر \mathbb{H} منتقل کرد $\Phi(x) \in \mathbb{H}$. در فضای جدید برای به دست آوردن ماتریس تبدیل برای استخراج $n(n < d)$ ویژگی و به دست آوردن $z = W^T x$ ، $z \in \mathbb{R}^n$ باید تابع هزینه زیر بیشینه شود

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

S_w^Φ و S_b^Φ به ترتیب ماتریس پراکنده‌گی بین کلاسی و درون کلاسی در فضای ویژگی \mathbb{H} هستند که به صورت زیر محاسبه می‌شوند

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

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

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 $\Phi(x_{i,k})$ i امین نمونه آموزشی در کلاس k ام است. Φ نشان‌دهنده این است که همه نمونه‌ها در فضای ویژگی \mathbb{H} هستند و محاسبات در این فضا انجام می‌شود. مشخص است که محاسبات نگاشت داده‌ها

ساختار محلی داده می‌باشد، شبیه به PCA برای کاهش ویژگی و بازسازی داده خوب کار می‌کنند و ممکن است برای کاربردهای طبقه‌بندی به اندازه کافی مطلوب نباشند. یک روش استخراج ویژگی نظارت شده دیگر نزدیک‌ترین نسبت دوتایی (DNP) 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 منفردشدن ماتریس پراکنده‌گی درون کلاسی استفاده می‌کند. روش دیگر استخراج ویژگی، عامل‌بندی ماتریس غیر منفی تبدیل متعامد با حفظ همسایگی (NPOPMMF) است [۱۶]. برای به دست آوردن ماتریس تبدیل NPOPMMF از دو عبارت در حل مسئله بهینه‌سازی آن استفاده می‌شود که یک عبارت مربوط به عامل‌بندی ماتریس غیر منفی تبدیل (PNMF) و عبارت دیگر مربوط به حفظ همسایگی محلی (NPO) است.

توجه کنید که هدف ما در این مقاله، استخراج ویژگی نظارت شده با استفاده از نمونه‌های آموزشی محدود است و در نتیجه روش پیشنهادی خود را با روش‌های نظارت شده استخراج ویژگی مقایسه می‌نماییم. لازم به ذکر است روش‌هایی چون NPE و LPP که هدف اصلی آنها حفظ ساختار محلی داده می‌باشد و به دو شیوه نظارت شده و بدون نظارت قابل پیاده‌سازی هستند، شبیه به PCA بیشتر به منظور کاهش ویژگی و بازسازی داده خوب استفاده شده و ممکن است برای کاربردهای طبقه‌بندی به اندازه کافی مطلوب نباشند. از میان روش‌های نظارت شده نیز معروف‌ترین روش یعنی LDA و بسط آن GDA را برای مقایسه انتخاب نموده‌ایم. همین طور روش غیر پارامتریک NWE که برای نمونه آموزشی کم، قابل اجرا می‌باشد را به عنوان یک روش معروف و کارا با روش پیشنهادی خود مقایسه کرده‌ایم. توجه کنید که سایر روش‌های نظارت شده‌ای که برای نمونه آموزشی کم مناسب هستند همانند DNP و NPOPMMF در تعداد نمونه‌های آموزشی محدود دارای نتایج مشابه NWE می‌باشند و لذا در انجام آزمایش‌های ما مورد ارزیابی و مقایسه قرار نمی‌گیرند.

در این مقاله یک روش استخراج ویژگی نظارت شده پیشنهاد گردیده که سعی می‌کند با استفاده از خوشه‌بندی، رابطه میان نمونه‌های آموزشی برچسب‌دار و نمونه‌های آزمایشی بدون برچسب را پیدا کند و سپس از قدرت نمونه‌های آزمایشی بدون برچسب برای تخمین ماتریس‌های پراکنده‌گی درون کلاسی و بین کلاسی استفاده نماید. روش پیشنهادی که تحلیل ممیز خطی فضای خوشه 1 (CSLDA) نامیده شده، دارای کارایی بسیار خوبی در تعداد نمونه‌های آموزشی محدود برای داده‌های ابرطیفی چندمنه می‌باشد. نتایج آزمایش‌های انجام شده بر روی تصاویر ابرطیفی واقعی، برتری روش CSLDA را نسبت به تعدادی از روش‌های استخراج ویژگی نظارت شده پرکاربرد نشان می‌دهد.

ادامه مقاله به صورت زیر سازمان یافته است: در بخش دوم مروری بر کارهای مرتبط با استخراج ویژگی خواهیم داشت. روش پیشنهادی در بخش سوم معرفی خواهد شد. نتایج حاصل از آزمایش‌ها و تحلیل آنها در بخش چهارم و در نهایت، نتیجه‌گیری در بخش پنجم بیان شده است.

۲- کارهای مرتبط

تحلیل ممیز خطی (LDA): یک روش کاهش ویژگی نظارت شده پارامتریک است که از معیار جدایی‌پذیری کلاس‌ها برای استخراج ویژگی

در نمونه‌های آموزشی و آزمایشی برای استخراج ویژگی استفاده کنیم. می‌دانیم که هر کدام از نمونه‌های آموزشی، متعلق به یکی از n_c کلاس موجود هستند. به مقداری که با بیشترین تکرار در یک مجموعه آماری رخ می‌دهد، مد گفته می‌شود. یک توزیع احتمالی چندمده^۱، یک توزیع احتمال با دو یا چند مد می‌باشد. هیستوگرام تصویر ابرطیفی چندمده دارای بیش از یک قله (مقدار بیشینه) می‌باشد. اگر داده ابرطیفی چندمده باشد، هر کدام از نمونه‌های آموزشی که برچسب یک کلاس خاص را خورده‌اند ممکن است به چندین خوشه متعلق باشند. بر این اساس ما بر روی داده آزمایشی، یک خوشه‌بندی با K خوشه انجام می‌دهیم. حال از خوشه‌بندی به دست آمده، ارتباط بین نمونه‌های آموزشی و آزمایشی را به دست آورده و از آن برای تولید ماتریس‌های پراکندگی درون کلاسی (S_w) و بین کلاسی (S_b) استفاده می‌کنیم

$$S_w = \sum_{c=1}^{n_c} \sum_{i=1}^{n_{ic}} (x_{ic} - m_c)(x_{ic} - m_c)^T + \quad (12)$$

$$\sum_{k=1}^K \sum_{x^k \in \mathcal{X}^k} p(k|x_{ic})(x^k - \mu_k)(x^k - \mu_k)^T$$

$$S_b = \sum_{c=1}^{n_c} n_{ic} (M_c - M)(M_c - M)^T \quad (13)$$

در روابط فوق x_{ic} نمونه آموزشی i ام از کلاس c ام، m_c میانگین نمونه‌های آموزشی کلاس c ام، n_{ic} تعداد نمونه‌های آموزشی کلاس c ام، x^k یک نمونه آزمایشی متعلق به خوشه k ام و \mathcal{X}^k مجموعه‌ای از نمونه‌های آزمایشی است که به خوشه k ام تعلق دارند. K تعداد خوشه‌ها و μ_k میانگین نمونه‌های آموزشی متعلق به خوشه k ام می‌باشد. $p(k|x_{ic})$ احتمال پسین تعلق نمونه آموزشی i ام از کلاس c ام (x_{ic}) به خوشه k ام است. میانگین ترکیبی کلاس c (M_c) و میانگین ترکیبی کل (M) که در S_b استفاده می‌شوند به صورت زیر محاسبه می‌گردند

$$M_c = m_c + \frac{1}{n_{ic}} \sum_{i=1}^{n_{ic}} \sum_{k=1}^K p(k|x_{ic}) \mu_k \quad (14)$$

$$M = \frac{1}{n_c} \sum_{c=1}^{n_c} m_c + \frac{1}{K} \sum_{k=1}^K \mu_k \quad (15)$$

همچنین میانگین نمونه‌های آموزشی کلاس c ام (m_c) و میانگین نمونه‌های آزمایشی خوشه k ام (μ_k) با استفاده از روابط زیر قابل محاسبه هستند

$$m_c = \frac{1}{n_{ic}} \sum_{i=1}^{n_{ic}} x_{ic} \quad (16)$$

$$\mu_k = \frac{1}{N_k} \sum_{x^k \in \mathcal{X}^k} x^k \quad (17)$$

در (۱۲) برای محاسبه S_w از دو قسمت استفاده شده است. در عبارت اول از نمونه‌های آموزشی و در عبارت دوم از نمونه‌های آزمایشی برای تولید ماتریس پراکندگی درون کلاسی استفاده شده است. در حقیقت با توجه به چندمده بودن داده، نمونه آموزشی i ام از کلاس c ام (x_{ic}) می‌تواند به چندین خوشه تعلق داشته باشد. احتمال تعلق x_{ic} به خوشه k ام ($k=1, 2, \dots, K$) به دست آمده و با توجه به میزان این احتمال، نمونه‌های آزمایشی موجود در آن خوشه در قسمت دوم عبارت S_w در

و سپس اجرای LDA از نظر محاسباتی پیچیده و زمان‌بر است. حتی ممکن است بعد فضای III بی‌نهایت باشد و به همین دلیل الگوریتم با جایگزینی ضرب‌های داخلی در فضای ویژگی جدید با تابع کرنل مناسب $k(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle$ بازنویسی می‌شود.

استخراج ویژگی وزن دار غیر پارامتریک (NWFE): این روش غیر پارامتریک بوده و نسبت به LDA برتری دارد. در حالی که LDA تنها قادر به استخراج $c-1$ ویژگی می‌باشد، روش NWFE با اختصاص وزن‌های متفاوت به هر نمونه برای محاسبه میانگین وزن دار و تعریف ماتریس‌های پراکندگی درون کلاسی و بین کلاسی غیر پارامتریک، قادر به استخراج بیش از $c-1$ ویژگی است. به علاوه NWFE با استفاده از تنظیم^۱ ماتریس پراکندگی درون کلاسی از منفرد شدن آن به هنگام استفاده از نمونه‌های آموزشی محدود، مقابله می‌کند. ماتریس‌های پراکندگی در این روش به شکل زیر تعریف شده‌اند

$$S_b = \sum_{i=1}^c p_i \sum_{j \neq i}^c \sum_{l=1}^{n_i} \frac{\lambda_l^{(i,j)}}{n_i} (x_l^{(i)} - M_j(x_l^{(i)}))(x_l^{(i)} - M_j(x_l^{(i)}))^T \quad (9)$$

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

در روابط فوق p_i احتمال پیشین کلاس i ام، $x_l^{(i)}$ نمونه l ام کلاس i و n_i تعداد نمونه آموزشی کلاس i ام است. از آنجایی که همه نمونه‌های آموزشی دارای اطلاعات یکسانی برای پیدا کردن مرز کلاس نیستند، NWFE با محاسبه میانگین وزن دار باعث بهبود نسبت به LDA می‌شود. 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

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

$$M_j(x_l^{(i)}) = \sum_{k=1}^{n_j} w_{lk}^{(i,j)} x_k^{(j)} \quad (10)$$

$$w_{lk}^{(i,j)} = \frac{\text{dist}(x_l^{(i)}, x_k^{(j)})^{-1}}{\sum_{i=1}^{n_j} \text{dist}(x_l^{(i)}, x_k^{(j)})^{-1}}$$

NWFE ماتریس S_w را به شیوه زیر تنظیم می‌کند

$$S_w = (1 - \alpha) S_w + \alpha \text{diag}(S_w) \quad (11)$$

در رابطه فوق $0 \leq \alpha \leq 1$ می‌باشد که $\alpha = 0.5$ انتخاب شده است. نتایج آزمایش‌ها بر روی داده‌های ابرطیفی نشان می‌دهند که انتخاب $\alpha = 0.5$ دقت طبقه‌بندی خوبی را فراهم می‌کند.

۳- تحلیل ممیز خطی فضای خوشه (CSLDA)

در شرایطی که تعداد نمونه‌های آموزشی در دسترس محدود است، تخمین ماتریس‌های پراکندگی درون کلاسی و بین کلاسی با دقت خوبی انجام نمی‌گیرد و با منفرد شدن ماتریس پراکندگی درون کلاسی، روش‌هایی چون LDA دارای کارایی خوبی نخواهند بود. ما در روش پیشنهادی خود سعی می‌کنیم که از هر دو اطلاعات موجود

ما از روش خوشه‌بندی پرکاربرد k-means به منظور خوشه‌بندی در این مقاله استفاده نموده‌ایم و تعداد خوشه‌ها را برابر تعداد کلاس‌ها ($K = n_c$) در نظر گرفته‌ایم. البته نوع خوشه‌بندی مورد استفاده و تعداد خوشه‌ها، انتخاب‌های آزادانه‌ای هستند. ما با توجه به نتایج مطلوبی که از داده‌های ابرطیفی مورد آزمایش به دست آورده‌ایم، روش k-means با $K = n_c$ را انتخاب نموده‌ایم.

یک ایراد روش CSLDA این است که رتبه ماتریس (تعداد سطرها یا ستون‌های مستقل خطی ماتریس) S_b در آن، برابر تعداد کلاس‌ها (n_c) می‌باشد که سبب می‌شود حداکثر بتوان n_c ویژگی با استفاده از این روش استخراج نمود. البته توجه کنید که رتبه ماتریس S_b در LDA و GDA، $n_c - 1$ است و این روش‌ها قادر به استخراج حداکثر $n_c - 1$ ویژگی هستند و در نتیجه، CSLDA نسبت به LDA و GDA قادر به استخراج یک ویژگی بیشتر است.

توجه داشته باشید که استفاده از روش CSLDA تنها برای تعداد نمونه آموزشی محدود توصیه می‌شود و هنگامی که نمونه آموزشی زیاد در دسترس است، چندان عاقلانه و به صرفه نیست چرا که زمانی که تعداد نمونه آموزشی به تعداد کافی و زیاد موجود است، استفاده از فقط نمونه‌های آموزشی به استفاده توأم از نمونه‌های آموزشی و آزمایشی ترجیح دارد و دقت طبقه‌بندی بهتری را فراهم می‌آورد. در ضمن با افزایش تعداد نمونه‌های آموزشی، پیچیدگی محاسباتی CSLDA افزایش یافته و زمان اجرای استخراج ویژگی داده را بالا می‌برد.

۴- پیاده‌سازی و نتایج آزمایش‌ها

۴-۱ داده‌ها و معیارهای ارزیابی مورد استفاده

در این مقاله از ۳ تصویر ابرطیفی برای انجام آزمایش‌ها استفاده شده است: ایندیانا، دانشگاه پاپوا و سالیاناس. ایندیانا تصویر ابرطیفی مربوط به منطقه‌ای کشاورزی- جنگلی است که توسط سنجنده AVIRIS از یک سایت آزمایشی در شمال شرقی ایالت ایندیانا گرفته شده است. این تصویر حاوی ۲۲۰ باند طیفی در بازه فرکانسی ۰/۴ تا ۲/۵ میکرومتر، ۱۴۵×۱۴۵ پیکسل و دقت مکانی ۲۰ متر بر پیکسل و ۱۶ کلاس است. بعد از حذف ۲۰ کانال نویزی، آزمایش‌ها بر روی ۲۰۰ باند باقیمانده انجام شده است. ما از میان ۱۶ کلاس داده ایندیانا از ۱۰ کلاس قابل توجه این داده که بیشترین تعداد نمونه‌های مرجع را دارا هستند، استفاده نموده‌ایم. طبقه‌بندی این داده به دلیل شرایط خاص منطقه با مشکلات زیادی مواجه است. یک 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 باقیمانده است. 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 دارند. همچنین بیشتر منطقه را زمین لخت یا خاشاک گیاهان سال‌های قبل پوشانده است. این عوامل باعث می‌شوند که توزیع احتمال کلاس‌ها چندمده باشند و در نتیجه، میزان جدایی‌پذیری کلاس‌ها کم و طبقه‌بندی داده دشوار خواهد بود. داده بعدی، تصویری شهری مربوط به محوطه دانشگاه پاپوا در ایتالیا است. این تصویر با استفاده از سنجنده نوری ROSIS اخذ شده و دارای تفکیک مکانی ۱/۳ متر بر پیکسل و دارای ۱۱۵ باند طیفی در بازه فرکانسی ۰/۴۳ تا ۰/۸۶ میکرومتر، ۹ کلاس و ابعاد ۶۱۰×۳۴۰ است. بعد از حذف کانال‌های نویزی،

تخمین ماتریس پراکندگی درون کلاسی مشارکت می‌کنند. همچنین M_c که میانگین ترکیبی کلاس c ام می‌باشد از دو قسمت تشکیل شده است. قسمت اول آن، میانگین نمونه‌های آموزشی کلاس c و قسمت دوم آن، میانگین نمونه‌های آزمایشی خوشه‌ها است. با توجه به میزان تعلق هر نمونه از کلاس c به خوشه k ام، میانگین آن خوشه در تولید M_c مشارکت خواهد داشت. M هم که میانگین کل ترکیبی است، از دو قسمت تشکیل شده که قسمت اول، میانگین نمونه‌های آموزشی و قسمت دوم، میانگین نمونه‌های آزمایشی می‌باشد.

با فرض معمول گوسی بودن داده ابرطیفی [۱۷] و [۱۸]، احتمال تعلق نمونه x_{ic} به خوشه k ام از رابطه زیر قابل محاسبه خواهد بود

$$p(x_{ic}|k) = \frac{1}{\sqrt{(2\pi)^d \Sigma_k}} \exp\left(-\frac{1}{2}(x_{ic} - \mu_k)^T \Sigma_k^{-1} (x_{ic} - \mu_k)\right) \quad (18)$$

در رابطه فوق، d تعداد ابعاد (ویژگی‌ها) داده ابرطیفی است. Σ_k و μ_k نیز به ترتیب بردار میانگین و ماتریس کواریانس خوشه k ام هستند. ماتریس Σ_k از رابطه زیر تخمین زده می‌شود

$$\Sigma_k = \frac{1}{N_k - 1} \sum_{x^k \in X^k} (x^k - \mu_k)(x^k - \mu_k)^T \quad (19)$$

N_k تعداد نمونه‌های آزمایشی متعلق به خوشه k ام می‌باشد. برای پرهیز از تخمین ماتریس کواریانس برای هر خوشه، فرض می‌کنیم که همه خوشه‌ها دارای توزیع کروی واحد هستند و ماتریس کواریانس آنها ماتریس همبندی (I) است. اگر تعداد کافی خوشه تولید شود، هر خوشه واریانس نسبتاً کوچکی دارد و این فرض عاقلانه به نظر می‌رسد. در نتیجه، تابع توزیع احتمال d بعدی برای خوشه k ام این گونه به دست می‌آید

$$p(x_{ic}|k) = \frac{1}{\sqrt{(2\pi)^d}} \exp\left(-\frac{1}{2} \text{dist}(x_{ic}, \mu_k)\right) \quad (20)$$

که $\text{dist}(x_{ic}, \mu_k) = (x_{ic} - \mu_k)^T (x_{ic} - \mu_k)$ فاصله اقلیدسی x_{ic} تا μ_k می‌باشد. احتمال پسین تعلق x_{ic} به خوشه k ام طبق (۲۱) خواهد بود

$$p(k|x_{ic}) = \frac{p(x_{ic}|k)p(k)}{\sum_{k=1}^K p(x_{ic}|k)p(k)} \quad (21)$$

در نتیجه با ادغام (۲۰) و (۲۱) خواهیم داشت

$$p(k|x_{ic}) = \frac{\exp\left(-\frac{1}{2} \text{dist}(x_{ic}, \mu_k)\right) p(k)}{\sum_{k=1}^K \exp\left(-\frac{1}{2} \text{dist}(x_{ic}, \mu_k)\right) p(k)} \quad (22)$$

در روابط فوق $p(k)$ احتمال پیشین خوشه k ام است که از جمعیت نسبی هر خوشه به دست می‌آید

$$p(k) = \frac{N_k}{N_{\text{test}}} \quad (23)$$

که N_{test} تعداد کل نمونه‌های آزمایشی می‌باشد.

ما همانند روش NWFEE از تنظیم ماتریس S_w برای جلوگیری از منفردشدن آن و افزایش دقت استفاده می‌کنیم. در انتها پس از محاسبه ماتریس‌های پراکندگی با بیشینه‌نمودن معیار $(tr(S_w^{-1} S_b))$ ماتریس تبدیل استخراج ویژگی به دست خواهد آمد.

یکی و با تابع هسته (کرنل) چندجمله‌ای درجه ۳ با مقادیر پیش‌فرض LIBSVM است.

معیارهای مختلفی برای ارزیابی کارایی طبقه‌بندها می‌توانند مورد استفاده قرار گیرند که از آن جمله می‌توان به دقت متوسط و اعتبار متوسط طبقه‌بندی اشاره کرد. دقت متوسط، میانگین دقت به دست آمده برای تمام کلاس‌ها و اعتبار متوسط نیز برابر میانگین اعتبار به دست آمده برای تمام کلاس‌ها می‌باشد. اعتبار برای هر کلاس این گونه تعریف شده است: تعداد نمونه‌هایی که درست طبقه‌بندی شده‌اند تقسیم بر تعداد کل نمونه‌هایی که به آن کلاس تعلق گرفته‌اند. ضریب کاپا یکی دیگر از پارامترهای دقت است که دقت طبقه‌بند را نسبت به یک طبقه‌بندی کاملاً تصادفی محاسبه می‌کند. ضریب کاپا به شکل زیر محاسبه می‌گردد [۲۰]

$$KC = \frac{N \sum_{c=1}^{n_c} t_{cc} - \sum_{c=1}^{n_c} t_{c+} t_{+c}}{N^2 - \sum_{c=1}^{n_c} t_{c+} t_{+c}} \quad (24)$$

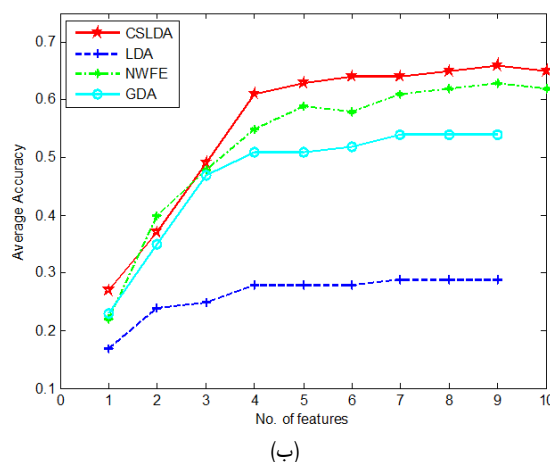
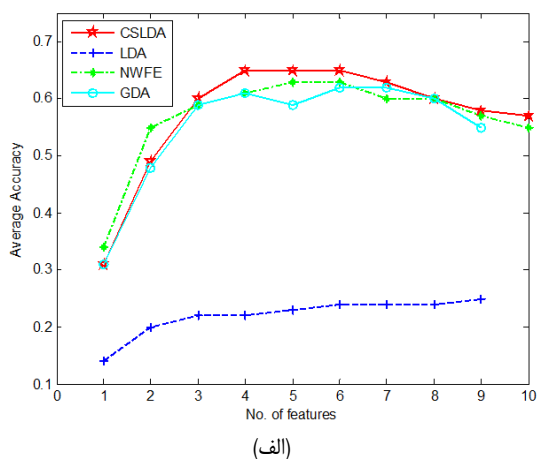
در رابطه فوق N و n_c به ترتیب تعداد نمونه‌های آزمایشی و تعداد کلاس‌ها هستند. t_{cc} تعداد نمونه‌هایی است که به طور صحیح در کلاس c طبقه‌بندی شده‌اند، t_{c+} تعداد نمونه‌های آزمایشی است که برچسب کلاس c را خورده‌اند و t_{+c} تعداد نمونه‌هایی است که پیشگویی می‌شود به کلاس c تعلق دارند. آزمون McNemars برای بیان معناداری تفاوت بین روش‌های طبقه‌بندی از لحاظ آماری بسیار مفید می‌باشد [۲۱]. این آزمون برای مقایسه بین یک جفت طبقه‌بند مورد استفاده قرار می‌گیرد. پارامتر Z_{12} به صورت زیر تعریف می‌شود

$$Z_{12} = \frac{f_{12} - f_{21}}{\sqrt{f_{12} + f_{21}}} \quad (25)$$

f_{12} تعداد نمونه‌هایی است که توسط طبقه‌بند ۱ درست طبقه‌بندی و توسط طبقه‌بند ۲ به اشتباه طبقه‌بندی شده‌اند. تفاوت بین طبقه‌بندهای ۱ و ۲ از نظر آماری با اهمیت گفته می‌شود اگر $|Z_{12}| > 1.96$. علامت پارامتر Z_{12} می‌گوید که طبقه‌بند ۱ دقیق‌تر از طبقه‌بند ۲ است ($Z_{12} > 0$) و بالعکس ($Z_{12} < 0$).

۴-۲ نتایج و تحلیل آزمایش‌ها

دقت متوسط طبقه‌بندی در برابر تعداد ویژگی‌های استخراج شده با استفاده از طبقه‌بندهای ML و SVM برای داده ایندینا در شکل ۱ رسم شده است. دقت و اعتبار به دست آمده برای تمام کلاس‌ها به همراه دقت متوسط طبقه‌بندی و ضریب کاپای به دست آمده با استفاده از طبقه‌بند SVM برای داده ایندینا با ۶ ویژگی استخراجی، داده پلویا با ۶ ویژگی استخراجی و داده سالیانس با ۱۰ ویژگی استخراجی، به ترتیب در جداول ۱ تا ۳ بیان شده‌اند. نقشه کلاس واقعی (GTM) و نقشه‌های کلاس به دست آمده نظیر این جدول‌ها نیز در شکل‌های ۲ تا ۴ نشان داده شده‌اند. مقادیر پارامتر Z حاصل از آزمون McNemars به ازای بالاترین دقت متوسط به دست آمده برای هر سه داده ابرطیفی و با استفاده از طبقه‌بندهای ML و SVM در جدول ۴ بیان گردیده‌اند. خلاصه مقادیر دقت طبقه‌بندی به دست آمده در جدول ۵ بیان شده‌اند. با توجه به نتایج به دست آمده از آزمایش‌ها، می‌توان به این نکات اشاره کرد: در هر سه داده ابرطیفی مورد آزمایش، روش LDA به علت منفرد شدن ماتریس پراکنندگی درون کلاسی، دارای کارایی بسیار ضعیفی در تعداد نمونه‌های آموزشی محدود می‌باشد.



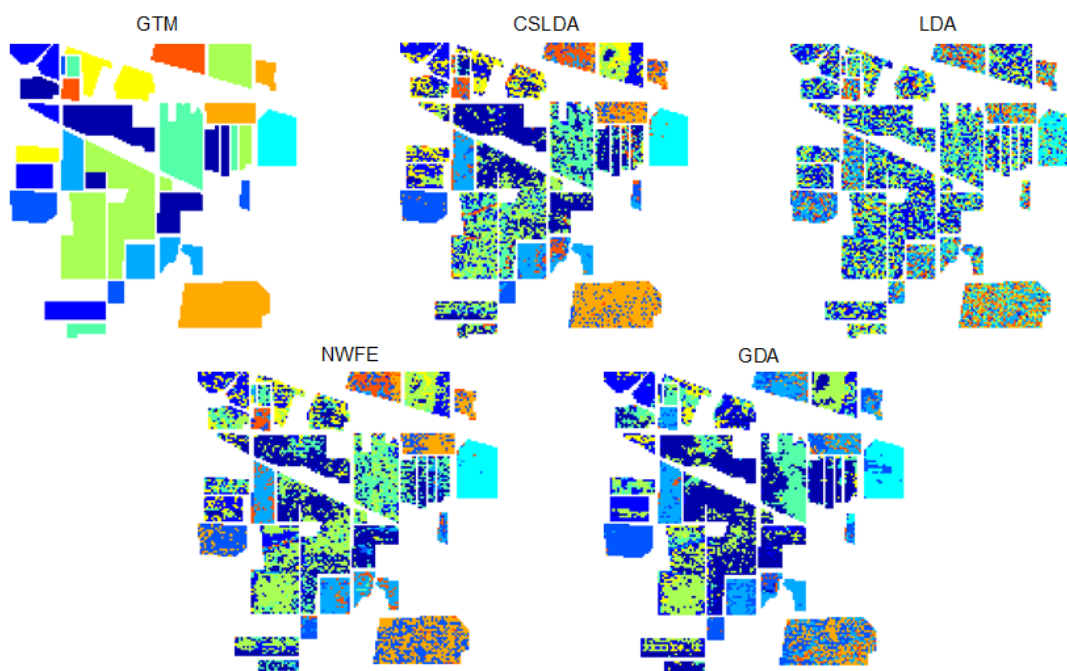
شکل ۱: دقت متوسط طبقه‌بندی در برابر تعداد ویژگی‌های استخراج شده برای داده ایندینا، (الف) با طبقه‌بند ML و (ب) با طبقه‌بند SVM.

آزمایش‌ها بر روی ۱۰۳ باند باقیمانده انجام شده است. داده سوم مورد استفاده، تصویر سالیانس است که از دهکده سالیانس در کالیفرنیا جنوبی توسط سنجنده AVIRIS اخذ شده است. این تصویر ۵۱۲×۲۱۷ پیکسل بوده و شامل ۱۶ کلاس، سایز پیکسل ۳/۷ متر و ۲۲۴ باند طیفی در بازه فرکانسی ۰/۴ تا ۲/۵ میکرومتر است. بعد از حذف ۲۰ باند جذب آب، ۲۰۴ باند از این تصویر در آزمایش‌ها استفاده می‌شوند.

ما برای بررسی کارایی روش‌های استخراج ویژگی در تعداد نمونه‌های آموزشی محدود، ۱۶ نمونه آموزشی را در هر کلاس در نظر گرفته و باقیمانده نمونه‌های هر کلاس را به عنوان نمونه آزمایشی استفاده می‌کنیم. نمونه‌های آموزشی به طور تصادفی از سراسر صحنه انتخاب می‌شوند. هر آزمایش ۱۰ مرتبه تکرار شده و متوسط نتایج به دست آمده گزارش شده است.

ما از دو طبقه‌بند برای ارزیابی روش‌های استخراج ویژگی استفاده نموده‌ایم. طبقه‌بند اول مورد استفاده، طبقه‌بند ماشین بردار پشتیبان (SVM) است. این طبقه‌بند غیر پارامتریک بوده و به تعداد نمونه‌های آموزشی، حساسیت کمتری دارد و به طور کلی برای تعداد نمونه‌های آموزشی محدود، مناسب و خوش‌رفتار است. طبقه‌بند دیگر مورد استفاده، طبقه‌بند بیشینه احتمال (ML) است. طبقه‌بند ML پارامتریک بوده و به تعداد نمونه‌های آموزشی بسیار حساس است.

ما برای استفاده از طبقه‌بند ML، توزیع معمول گوسی را برای کلاس‌ها در نظر گرفته‌ایم. همچنین برای پیاده‌سازی طبقه‌بند SVM از کتابخانه LIBSVM بهره برده‌ایم [۱۹]. طبقه‌بند SVM مورد استفاده در انجام آزمایش‌ها در حالت چند کلاسه با استفاده از روش یکی در برابر



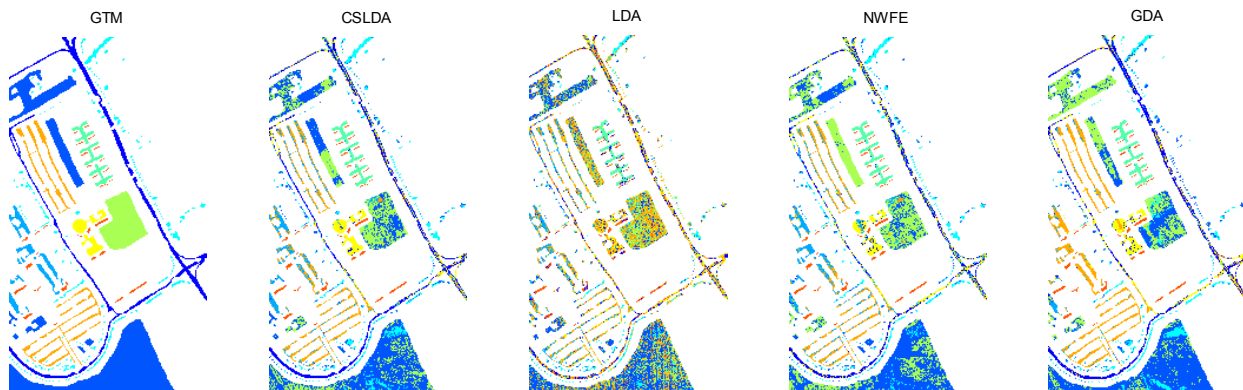
شکل ۲: نقشه‌های کلاس به دست آمده برای داده ایندیانا با استفاده از ۶ ویژگی استخراجی و طبقه‌بند SVM.

جدول ۱: دقت و اعتبار کلاس‌ها برای داده ایندیانا با استفاده از ۶ ویژگی استخراجی و طبقه‌بند SVM.

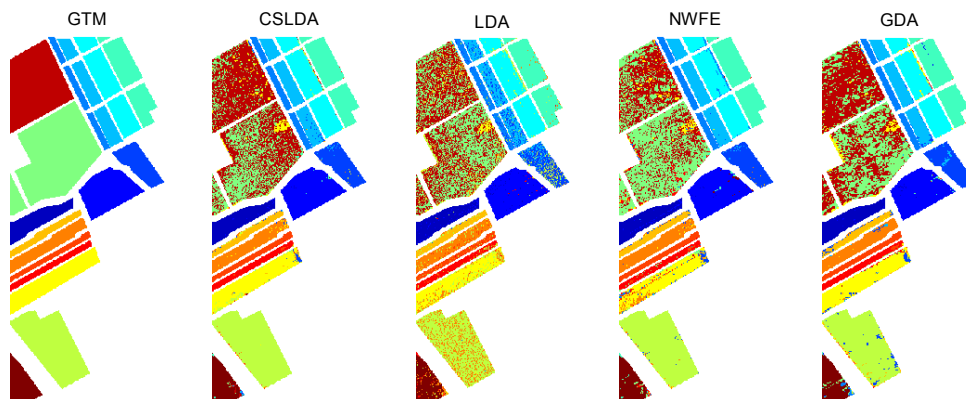
شماره	کلاس	نام کلاس	تعداد نمونه‌ها	CSLDA		LDA		NWFE		GDA	
				دقت	اعتبار	دقت	اعتبار	دقت	اعتبار	دقت	اعتبار
۱	ذرت- بدون شخم	۱۴۳۴	۰٫۷۷	۰٫۴۸	۰٫۱۷	۰٫۲۸	۰٫۶۳	۰٫۴۳	۰٫۷۲	۰٫۳۷	
۲	ذرت- کم شخم	۸۳۴	۰٫۴۱	۰٫۳۹	۰٫۳۵	۰٫۱۶	۰٫۴۲	۰٫۴۴	۰٫۵۹	۰٫۴۶	
۳	سیزه- درختان	۴۹۷	۰٫۸۷	۰٫۵۱	۰٫۲۵	۰٫۳۳	۰٫۷۰	۰٫۳۴	۰٫۸۹	۰٫۳۶	
۴	سیزه- چمنزار	۷۴۷	۰٫۷۳	۰٫۸۷	۰٫۳۱	۰٫۱۶	۰٫۷۷	۰٫۷۳	۰٫۷۳	۰٫۴۵	
۵	کاه و خاشاک	۴۸۹	۰٫۹۸	۱٫۰۰	۰٫۵۰	۰٫۷۴	۰٫۹۸	۱٫۰۰	۰٫۸۹	۰٫۹۳	
۶	سویا- بدون شخم	۹۶۸	۰٫۴۷	۰٫۴۲	۰٫۳۹	۰٫۱۷	۰٫۴۷	۰٫۳۸	۰٫۴۹	۰٫۵۱	
۷	سویا- کم شخم	۲۴۶۸	۰٫۳۱	۰٫۶۵	۰٫۱۱	۰٫۴۱	۰٫۴۳	۰٫۶۷	۰٫۳۳	۰٫۷۱	
۸	سویا- شخم کامل	۶۱۴	۰٫۴۵	۰٫۴۱	۰٫۱۸	۰٫۱۵	۰٫۳۱	۰٫۴۲	۰٫۱۷	۰٫۳۲	
۹	بیشه	۱۲۹۴	۰٫۷۴	۰٫۹۵	۰٫۲۶	۰٫۵۶	۰٫۵۵	۰٫۸۳	۰٫۲۹	۰٫۹۵	
۱۰	ساختمان- سیزه	۳۸۰	۰٫۶۴	۰٫۳۹	۰٫۲۹	۰٫۱۷	۰٫۵۱	۰٫۴۴	۰٫۰۶	۰٫۱۴	
دقت متوسط و اعتبار متوسط			۰٫۶۴	۰٫۶۱	۰٫۲۸	۰٫۳۱	۰٫۵۸	۰٫۵۷	۰٫۵۲	۰٫۵۲	
ضریب کاپا			۰٫۵۲	۰٫۱۶	۰٫۴۸	۰٫۴۲	۰٫۴۸	۰٫۴۲	۰٫۴۲	۰٫۴۲	

جدول ۲: دقت و اعتبار کلاس‌ها برای داده پلویا با استفاده از ۶ ویژگی استخراجی و طبقه‌بند SVM.

شماره	کلاس	نام کلاس	تعداد نمونه‌ها	CSLDA		LDA		NWFE		GDA	
				دقت	اعتبار	دقت	اعتبار	دقت	اعتبار	دقت	اعتبار
۱	آسفالت	۶۶۳۱	۰٫۶۱	۰٫۹۷	۰٫۲۸	۰٫۶۵	۰٫۴۷	۰٫۹۳	۰٫۵۹	۰٫۹۵	
۲	چمنزار	۱۸۶۴۹	۰٫۶۶	۰٫۸۱	۰٫۵۲	۰٫۸۸	۰٫۵۸	۰٫۸۲	۰٫۶۵	۰٫۷۸	
۳	شن و سنگریزه	۲۰۹۹	۰٫۶۵	۰٫۴۳	۰٫۴۸	۰٫۳۲	۰٫۷۴	۰٫۴۲	۰٫۲۰	۰٫۴۰	
۴	درختان	۳۰۶۴	۰٫۶۸	۰٫۶۷	۰٫۴۳	۰٫۶۸	۰٫۷۲	۰٫۸۰	۰٫۵۸	۰٫۵۷	
۵	صفحات فلزی	۱۳۴۵	۰٫۹۸	۱٫۰۰	۰٫۸۴	۱٫۰۰	۰٫۹۸	۰٫۹۸	۰٫۹۹	۰٫۸۴	
۶	زمین دست‌نخورده	۵۰۲۹	۰٫۵۹	۰٫۳۵	۰٫۳۲	۰٫۲۳	۰٫۶۵	۰٫۳۰	۰٫۴۳	۰٫۲۷	
۷	قییر	۱۳۳۰	۰٫۹۱	۰٫۴۵	۰٫۳۴	۰٫۱۸	۰٫۷۴	۰٫۳۰	۰٫۷۸	۰٫۳۵	
۸	آجر	۳۶۸۲	۰٫۷۴	۰٫۷۱	۰٫۴۶	۰٫۱۴	۰٫۷۳	۰٫۷۹	۰٫۹۰	۰٫۶۰	
۹	سایه	۹۴۷	۱٫۰۰	۱٫۰۰	۰٫۷۹	۰٫۴۳	۰٫۹۹	۰٫۹۸	۱٫۰۰	۱٫۰۰	
دقت متوسط و اعتبار متوسط			۰٫۷۶	۰٫۷۱	۰٫۴۹	۰٫۵۰	۰٫۷۳	۰٫۷۰	۰٫۶۸	۰٫۶۴	
ضریب کاپا			۰٫۵۹	۰٫۳۴	۰٫۵۴	۰٫۵۴	۰٫۵۴	۰٫۵۴	۰٫۵۳	۰٫۵۳	



شکل ۳: نقشه‌های کلاس به دست آمده برای داده پاپویا با استفاده از ۶ ویژگی استخراجی و طبقه‌بند SVM.



شکل ۴: نقشه‌های کلاس به دست آمده برای داده سالیناس با استفاده از ۱۰ ویژگی استخراجی و طبقه‌بند SVM.

جدول ۳: دقت و اعتبار کلاس‌ها برای داده سالیناس با استفاده از ۱۰ ویژگی استخراجی و طبقه‌بند SVM.

شماره	نام کلاس	تعداد نمونه‌ها	CSLDA		LDA		NWFE		GDA	
			دقت	اعتبار	دقت	اعتبار	دقت	اعتبار	دقت	اعتبار
۱	بروکلی- سبزه- علف هرز ۱	۲۰۰۹	۰.۹۸	۰.۹۹	۰.۹۱	۰.۹۵	۰.۹۳	۰.۹۹	۰.۹۸	۰.۹۹
۲	بروکلی- سبزه- علف هرز ۲	۳۷۲۶	۰.۹۹	۰.۹۹	۰.۹۶	۰.۹۶	۰.۹۷	۰.۹۶	۰.۹۸	۰.۹۹
۳	زمین شخم‌شده	۱۹۷۶	۰.۹۷	۰.۹۰	۰.۶۴	۰.۶۴	۰.۹۷	۰.۸۸	۰.۹۵	۰.۷۷
۴	زمین شخم‌شده- زبر- گاواهن	۱۳۹۴	۰.۹۹	۰.۹۸	۰.۸۵	۰.۹۱	۱.۰۰	۰.۷۵	۱.۰۰	۰.۹۱
۵	زمین شخم‌شده- هموار	۲۶۷۸	۰.۹۷	۰.۹۷	۰.۷۴	۰.۸۵	۰.۹۸	۰.۹۸	۰.۹۵	۰.۹۳
۶	کاه بن	۳۹۵۹	۰.۹۶	۱.۰۰	۰.۹۲	۱.۰۰	۰.۹۴	۱.۰۰	۰.۹۵	۱.۰۰
۷	کرفس	۳۵۷۹	۱.۰۰	۰.۹۸	۰.۹۷	۱.۰۰	۰.۹۹	۰.۹۷	۰.۹۹	۰.۹۷
۸	انگور	۱۱۲۷۱	۰.۴۷	۰.۸۰	۰.۴۹	۰.۷۰	۰.۶۸	۰.۷۲	۰.۵۱	۰.۷۸
۹	خاک- تاکستان	۶۲۰۳	۰.۹۹	۰.۹۶	۰.۷۵	۰.۶۲	۰.۹۸	۰.۹۹	۰.۹۳	۰.۹۹
۱۰	ذرت- سبزه پیر- علف هرز	۳۲۷۸	۰.۸۹	۰.۸۰	۰.۷۱	۰.۸۷	۰.۷۵	۰.۷۹	۰.۸۹	۰.۷۶
۱۱	کاهو- ۴ هفته	۱۰۶۸	۰.۸۵	۰.۹۰	۰.۸۵	۰.۹۴	۰.۸۸	۰.۹۶	۰.۸۱	۰.۸۱
۱۲	کاهو- ۵ هفته	۱۹۲۷	۰.۹۵	۰.۹۶	۰.۸۶	۰.۳۶	۰.۹۷	۰.۸۴	۰.۹۶	۰.۹۷
۱۳	کاهو- ۶ هفته	۹۱۶	۰.۹۹	۰.۹۱	۰.۸۱	۰.۷۵	۰.۹۸	۰.۸۰	۱.۰۰	۰.۹۲
۱۴	کاهو- ۷ هفته	۱۰۷۰	۰.۹۵	۰.۹۳	۰.۸۳	۰.۸۸	۰.۸۷	۰.۵۶	۰.۹۰	۰.۸۵
۱۵	تاکستان	۷۲۶۸	۰.۸۰	۰.۵۱	۰.۵۷	۰.۴۷	۰.۵۶	۰.۵۹	۰.۷۷	۰.۵۳
۱۶	تاکستان- شاغولی- داریست	۱۸۰۷	۰.۹۲	۰.۹۹	۰.۸۴	۱.۰۰	۰.۹۰	۰.۹۶	۰.۸۹	۱.۰۰
دقت متوسط و اعتبار متوسط			۰.۹۲	۰.۹۱	۰.۷۹	۰.۸۱	۰.۹۰	۰.۸۶	۰.۹۰	۰.۸۹
ضریب کاپا			۰.۸۲		۰.۷۰		۰.۸۲		۰.۸۱	

دقت متوسط طبقه‌بندی به دست آمده با روش CSLDA نسبت به روش‌های NWFE و GDA به ترتیب ۶٪ و ۱۲٪ افزایش یافته است. این افزایش دقت CSLDA نسبت به NWFE و GDA به ازای ۴ ویژگی استخراجی و طبقه‌بند ML، ۴٪ بوده است. برای داده پاپویا که از نظر سختی طبقه‌بندی می‌توان آن را بین داده

با توجه به این که داده ایندیانا، داده‌ای چندمنده و سخت برای طبقه‌بندی است، استفاده از یک روش استخراج ویژگی خوب و مناسب، دقت طبقه‌بندی را در آن به میزان قابل ملاحظه‌ای افزایش می‌دهد و تفاوت بین روش‌های مناسب و نامناسب در این داده بیشتر آشکار می‌گردد. به عنوان نمونه به ازای ۶ ویژگی استخراجی و طبقه‌بند SVM،

جدول ۴: اهمیت آماری تفاوت در دقت بین طبقه‌بندها با استفاده از معیار آزمون MCNEMARS (هر خانه جدول بیان‌کننده Z_{RC} است که R شماره سطر و C شماره ستون جدول است).

داده ایندیانا، طبقه‌بند SVM، ۶ ویژگی استخراجی					داده ایندیانا، طبقه‌بند ML، ۴ ویژگی استخراجی				
	CSLDA	LDA	NWFE	GDA		CSLDA	LDA	NWFE	GDA
CSLDA	۰	۴۶٫۹۵	۶٫۲۹	۱۵٫۸۷	CSLDA	۰	۵۲٫۶۴	۶٫۴۴	۴٫۰۲
LDA	-۴۶٫۹۵	۰	-۴۲٫۴۰	-۳۵٫۸۵	LDA	-۵۲٫۶۴	۰	-۴۹٫۸۳	-۵۰٫۶۰
NWFE	-۶٫۲۹	۴۲٫۴۰	۰	۹٫۱۳	NWFE	-۶٫۴۴	۴۹٫۸۳	۰	-۱٫۸۶
GDA	-۱۵٫۸۷	۳۵٫۸۵	-۹٫۱۳	۰	GDA	-۴٫۰۲	۵۰٫۶۰	۱٫۸۶	۰

داده پالویا، طبقه‌بند SVM، ۶ ویژگی استخراجی					داده پالویا، طبقه‌بند ML، ۴ ویژگی استخراجی				
	CSLDA	LDA	NWFE	GDA		CSLDA	LDA	NWFE	GDA
CSLDA	۰	۶۶٫۵۸	۱۸٫۰۲	۱۵٫۲۱	CSLDA	۰	۷۰٫۳۳	۲۷٫۱۴	-۲٫۸۹
LDA	-۶۶٫۵۸	۰	-۵۲٫۸۵	-۵۲٫۸۷	LDA	-۷۰٫۳۳	۰	-۵۷٫۸۷	-۷۴٫۱۰
NWFE	-۱۸٫۰۲	۵۲٫۸۵	۰	-۱٫۱۵	NWFE	-۲۷٫۱۴	۵۷٫۸۷	۰	-۲۱٫۸۵
GDA	-۱۵٫۲۱	۵۲٫۸۷	۱٫۱۵	۰	GDA	۲٫۸۹	۷۴٫۱۰	۲۱٫۸۵	۰

داده سالیانس، طبقه‌بند SVM، ۱۰ ویژگی استخراجی					داده سالیانس، طبقه‌بند ML، ۶ ویژگی استخراجی				
	CSLDA	LDA	NWFE	GDA		CSLDA	LDA	NWFE	GDA
CSLDA	۰	۵۲٫۲۶	۲٫۲۱	۴٫۱۷	CSLDA	۰	۶۱٫۳۲	۵٫۷۸	-۴٫۳۵
LDA	-۵۲٫۲۶	۰	-۴۹٫۵۱	-۵۰٫۱۲	LDA	-۶۱٫۳۲	۰	-۵۷٫۸۲	-۶۳٫۹۲
NWFE	-۲٫۲۱	۴۹٫۵۱	۰	۱٫۸۶	NWFE	-۵٫۷۸	۵۷٫۸۲	۰	-۹٫۷۴
GDA	-۴٫۱۷	۵۰٫۱۲	-۱٫۸۶	۰	GDA	۴٫۳۵	۶۳٫۹۲	۹٫۷۴	۰

GDA، ۳٪ کاهش دقت داشته است. چون داده سالیانس، چندمده نیست، روش CSLDA بر روی آن چندان خوب عمل نمی‌کند. مشاهده می‌کنیم که نتایج آزمایش‌ها مطابق با انتظار ما است. همان‌طور که قبلاً هم گفتیم، روش استخراج ویژگی CSLDA برای داده‌های ابرطیفی چندمده دارای کارایی خوبی بوده و نسبت به سایر روش‌ها، برتری دارد.

با توجه به نتایج به دست آمده می‌توان مشاهده کرد که به طور کلی تفاوت بین روش‌ها با استفاده از طبقه‌بند SVM بیشتر و با استفاده از طبقه‌بند ML کمتر است. علت این امر را می‌توان قدرت بالای طبقه‌بند SVM دانست که در تعداد نمونه‌های آموزشی محدود، دارای کارایی خوبی است، در حالی که کارایی طبقه‌بند ML به شدت به تعداد نمونه‌های آموزشی بستگی دارد. به گونه‌ای که برای داشتن یک طبقه‌بندی خوب با ML، تعداد نمونه‌های آموزشی حدوداً باید ۱۰ برابر تعداد ویژگی‌ها باشد که در شرایطی که با محدودیت نمونه‌های آموزشی مواجه‌ایم، این امر مقدر نیست.

۵- نتیجه‌گیری

در این مقاله یک روش استخراج ویژگی نظارت‌شده برای داده‌های ابرطیفی معرفی شد. روش پیشنهادی که CSLDA نام دارد دارای کارایی خوبی برای تصاویر ابرطیفی چندمده با استفاده از تعداد نمونه‌های آموزشی محدود می‌باشد. روش CSLDA داده آزمایشی را خوشه‌بندی کرده و سپس با یافتن رابطه بین نمونه‌های آزمایشی و آموزشی، از قدرت نمونه‌های بدون برچسب آزمایشی به همراه تعداد محدود نمونه‌های آموزشی در دسترس، برای تخمین ماتریس پراکنندگی درون‌کلاسی و بین‌کلاسی استفاده می‌کند. نتایج آزمایش‌های انجام‌شده بر روی تصاویر ابرطیفی واقعی نشان می‌دهند که CSLDA نسبت به روش‌های پرکاربرد استخراج ویژگی نظارت‌شده در شرایطی که نمونه‌های آموزشی کمی در دسترس هستند، برتری دارد.

جدول ۵: بالاترین دقت طبقه‌بندی به دست آمده برای روش‌های مختلف استخراج ویژگی (عدد نوشته‌شده در پرانتزها، تعداد ویژگی‌هایی را که بالاترین دقت متوسط طبقه‌بندی را فراهم کرده‌اند، نشان می‌دهد. در هر آزمایش، اعداد مربوط به روشی که بالاترین دقت را فراهم کرده‌اند پررنگ شده‌اند).

	داده	طبقه‌بند	CSLDA	LDA	NWFE	GDA
ایندیانا	SVM	۰٫۶۶	۰٫۲۹	۰٫۶۴	۰٫۵۴	
		(۹)	(۷)	(۱۱)	(۷)	
پالویا	ML	۰٫۶۵	۰٫۲۵	۰٫۶۳	۰٫۶۲	
		(۴)	(۹)	(۵)	(۶)	
پالویا	SVM	۰٫۷۷	۰٫۵۱	۰٫۷۷	۰٫۷۴	
		(۹)	(۸)	(۸)	(۷)	
سالیانس	ML	۰٫۸۰	۰٫۴۸	۰٫۷۸	۰٫۸۰	
		(۴)	(۷)	(۶)	(۷)	
سالیانس	SVM	۰٫۹۲	۰٫۸۰	۰٫۹۰	۰٫۹۰	
		(۹)	(۱۱)	(۹)	(۵)	
سالیانس	ML	۰٫۸۸	۰٫۷۸	۰٫۹۰	۰٫۹۲	
		(۶)	(۸)	(۴)	(۳)	

ایندیانا و سالیانس در نظر گرفت، با استفاده از ۶ ویژگی استخراجی و طبقه‌بند SVM، CSLDA نسبت به NWFE و GDA به ترتیب ۳٪ و ۸٪ بهبود دقت داشته است. همچنین با ۴ ویژگی استخراجی و طبقه‌بند ML، CSLDA نسبت به هر دو روش NWFE و GDA، ۳٪ افزایش دقت دارد.

داده سالیانس از نظر طبقه‌بندی نسبت به دو داده دیگر ساده‌تر بوده و چندمده نمی‌باشد، به همین دلیل تفاوت بین روش‌های استخراج ویژگی در طبقه‌بندی این داده، چندان ملموس نیست. به عنوان مثال با ۱۰ ویژگی استخراجی و طبقه‌بند SVM، CSLDA نسبت به هر دو روش NWFE و GDA دارای ۲٪ افزایش دقت است. به ازای ۶ ویژگی استخراجی و طبقه‌بند ML، CSLDA نسبت به NWFE، ۱٪ و نسبت به

مراجع

- [13] X. F. He and P. Niyogi, "Locality preserving projections," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 16, pp. 153-160, 2004.
- [14] X. F. He, D. Cai, S. C. Yan, and H. J. Zhang, "Neighborhood preserving embedding," in *Proc. 10th IEEE Int. Conf. Comput. Vis., ICCV'05*, vol. 2, pp. 1208-1213, 17-21 Oct. 2005.
- [15] H. Y. Huang and B. C. Kuo, "Double nearest proportion feature extraction for hyperspectral-image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 11, pp. 4034-4046, Nov. 2010.
- [16] J. Wen, Z. Tian, X. Liu, and W. Lin, "Neighborhood preserving orthogonal PNMf feature extraction for hyperspectral image classification," *IEEE J. of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 2, pp. 759-768, Apr. 2013.
- [17] A. M. MartoAnez and A. C. Kak, "PCA versus LDA," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, no. 2, pp. 228-233, Feb. 2001.
- [18] J. H. Friedman, "Regularized discriminant analysis," *J. of the American Statistical Association*, vol. 84, no. 405, pp. 165-175, Mar. 1989.
- [19] C. Chang and C. Lin, LIBSVM-A Library for Support Vector Machines, 2008. [Online], Available: <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [20] J. Cohen, "A coefficient of agreement from nominal scales," *Edu. Psychol. Meas.*, vol. 20, no. 1, pp. 37-46, Apr. 1960.
- [21] G. M. Foody, "Thematic map comparison: evaluating the statistical significance of differences in classification accuracy," *Photogramm. Eng. Remote Sens.*, vol. 70, no. 5, pp. 627-633, May 2004.
- [1] G. F. Hughes, "On the mean accuracy of statistical pattern recognition," *IEEE Trans. Inf. Theory*, vol. 14, no. 1, pp. 55-63, Jan. 1968.
- [2] G. Moser and S. B. Serpico, "Combining support vector machines and markov random fields in an integrated framework for contextual image classification," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 51, no. 5, pp. 2734-2752, May 2013.
- [3] B. Demir and S. Erturk, "Clustering-based extraction of border training patterns for accurate SVM classification of hyperspectral images," *IEEE Geoscience and Remote Sensing Lett.*, vol. 6, no. 4, pp. 840-844, Oct. 2009.
- [4] M. Imani and H. Ghassemian, "Adaptive expansion of training samples for improving hyperspectral image classification performance," in *Proc. 21st Iranian Conf. on Electrical Engineering*, 6 pp., May 2013.
- [5] X. Jia, B. C. Kuo, and M. Crawford, "Feature mining for hyperspectral image classification," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 676-697, Mar. 2013.
- [6] M. Kamandar and H. Ghassemian, "Linear feature extraction for hyperspectral images based on information theoretic learning," *IEEE Geoscience and Remote Sensing Lett.*, vol. 10, no. 4, pp. 702-706, Jun. 2013.
- [7] M. Imani and H. Ghassemian, "Band clustering-based feature extraction for classification of hyperspectral images using limited training samples," *IEEE Geoscience and Remote Sensing Lett.*, vol. 11, no. 8, pp. 1325-1329, Aug. 2014.
- [8] D. Korycinski, M. Crawford, J. W. Barnes, and J. Ghosh, "Adaptive feature selection for hyperspectral data analysis using a binary hierarchical classifier and tabu search," in *Proc. IEEE Symp. on Geoscience and Remote Sensing*, vol. 1, pp. 297-299, Jul. 2003.
- [9] S. Li, H. Wu, D. Wan, and J. Zhu, "An effective feature selection method for hyperspectral image classification based on genetic algorithm and support vector machine," *Knowledge-Based Systems*, vol. 24, no. 1, pp. 40-48, Feb. 2011.
- [10] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, San Diego: Academic Press Inc., 1990.
- [11] G. Baudat and F. Anouar, "Generalized discriminant analysis using a kernel approach," *Neural Comput.*, vol. 12, no. 10, pp. 2385-2404, Oct. 2000.
- [12] B. C. Kuo and D. A. Landgrebe, "Nonparametric weighted feature extraction for classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 5, pp. 1096-1105, May 2004.

مریم ایمانی تحصیلات کارشناسی و کارشناسی ارشد خود را در رشته مهندسی برق - گرایش مخابرات به ترتیب در سال‌های ۱۳۸۸ و ۱۳۹۰ در دانشگاه شاهد به پایان رساند. ایشان در سال ۱۳۹۴ مدرک دکتری خود را در همان رشته از دانشگاه تربیت مدرس اخذ نمود. زمینه‌های تحقیقاتی مورد علاقه ایشان، شناسایی آماری الگو، پردازش اطلاعات (سیگنال و تصویر) و مهندسی سنجش از دور می‌باشد.

حسن قاسمیان تحصیلات کارشناسی خود را در رشته مهندسی مخابرات در سال ۱۳۵۸ در دانشکده مخابرات تهران به پایان رساند. ایشان مدرک کارشناسی ارشد و دکتری خود را در رشته مهندسی مخابرات در دانشگاه پردو آمریکا به ترتیب در سال ۱۳۶۳ و ۱۳۶۷ اخذ کرد. ایشان هم‌اکنون استاد دانشکده مهندسی برق و کامپیوتر هستند. آنالیز و پردازش تصاویر از منابع چندگانه، پردازش اطلاعات و شناسایی الگو، مهندسی سنجش از دور، پردازش تصاویر و سیگنال‌های مهندسی پزشکی از جمله علاقه‌های پژوهشی ایشان است.