

جاسازی خط ویژگی وزن دار برای استخراج ویژگی تصاویر ابرطیفی

مریم ایمانی^۱ حسن قاسمیان^۲

۱- دانش‌آموخته کارشناسی ارشد- دانشکده مهندسی برق و کامپیوتر- دانشگاه تربیت مدرس- تهران- ایران
maryam.imani@modares.ac.ir

۲- استاد دانشکده مهندسی برق و کامپیوتر- دانشگاه تربیت مدرس- تهران- ایران
ghassemi@modares.ac.ir

چکیده: یکی از مراحل مهم قبل از طبقه‌بندی تصاویر ابرطیفی، کاهش ویژگی با استفاده از روش‌های استخراج ویژگی است. در این مقاله یک روش استخراج نظارت شده پیشنهاد شده که دارای کارایی خوبی با استفاده از تعداد نمونه‌های آموزشی محدود است. روش استخراج ویژگی پیشنهادی که جاسازی خط ویژگی وزن دار (WFLE) نامیده شده، از مفاهیم خط ویژگی برای تولید نمونه‌های آموزشی مجازی استفاده می‌کند. نمونه‌های آموزشی مجازی تولید شده برای محاسبه ماتریس‌های پراکندگی درون دسته‌ای و بین دسته‌ای به شکل وزن دار استفاده می‌شوند. نحوه وزن‌دهی بر مبنای ماهیت نمونه‌های آموزشی است. آن دسته از نمونه‌های آموزشی که سبب ایجاد خطا در طبقه‌بندی داده می‌شوند، نمونه‌های آموزشی نامطلوب محسوب شده و بنابراین در طی فرآیند استخراج ویژگی اصلاح بیشتری بر روی آن‌ها انجام می‌شود. در مقابل، بر روی نمونه‌های آموزشی مطلوب، اصلاح کمتری صورت می‌پذیرد. روش پیشنهادی WFLE با تعدادی از روش‌های استخراج ویژگی مهم و پرکاربرد از قبیل LDA، GDA، NWFE، LPP، NPE و NFLE مقایسه شده است. ما برای انجام آزمایش‌های خود از سه مجموعه داده ابرطیفی واقعی استفاده کرده‌ایم. نتایج آزمایش‌ها، برتری روش پیشنهادی را نسبت به سایر روش‌ها در تعداد نمونه‌های آموزشی محدود نشان می‌دهند.

کلمات کلیدی: خط ویژگی، استخراج ویژگی، تصویر ابرطیفی، طبقه‌بندی

تاریخ ارسال مقاله: ۱۳۹۳/۱/۲۳

تاریخ پذیرش مقاله: ۱۳۹۴/۹/۱۳

نام نویسنده‌ی مسئول: دکتر حسن قاسمیان

نشانی نویسنده‌ی مسئول: بزرگراه جلال آل احمد، پل نصر، دانشگاه تربیت مدرس، دانشکده مهندسی برق و کامپیوتر



انجام گیرد و برعکس، اگر نمونه‌های آموزشی دارای وضعیت عادی و مطلوبی باشند، اصلاح کمتری بر روی آن‌ها انجام خواهد گرفت.

ما روش استخراج ویژگی پیشنهادی، WFLE، را با سه گروه از روش‌های استخراج ویژگی مقایسه کرده‌ایم. گروه اول، روش‌های استخراج ویژگی‌ای هستند که بر مبنای جدایی‌پذیری دسته‌ها کار می‌کنند. این روش‌ها با تولید ماتریس‌های پراکندگی درون دسته‌ای و بین دسته‌ای و بیشینه کردن پراکندگی‌های بین دسته‌ای دارد. از جمله‌ی این روش‌ها می‌توان به تحلیل تمیز خطی^(۴) (LDA) [۱۰]، تحلیل تمیز تعمیم یافته^(۵) (GDA) [۱۱] و استخراج ویژگی وزن‌دار غیرپارامتریک^(۶) (NWFE) [۱۲] اشاره نمود. گروه دوم روش‌های استخراج ویژگی، به منظور حفظ ساختار محلی داده استفاده می‌شوند. از جمله این روش‌ها می‌توان به جاسازی حفظ هماسیگی^(۷) (NPE) [۱۳] و تصویر حفظ محلیت^(۸) (LPP) [۱۴] اشاره نمود که می‌توانند به دو صورت نظارت شده و بدون نظارت، پیاده‌سازی شوند که البته در این مقاله، فرم نظارت شده‌ی آن‌ها برای ما مد نظر است. و اما گروه سوم روش‌های استخراج ویژگی که در این‌جا بررسی می‌شود، روش‌هایی هستند که از نمونه‌های آموزشی مجازی استفاده می‌کنند. از جمله این روش‌ها می‌توان به NFLE^(۹) [۱۵] اشاره نمود که از نمونه‌های آموزشی مجازی تولید شده برای تخمین ماتریس‌های پراکندگی درون دسته‌ای و بین دسته‌ای استفاده می‌کند.

در ادامه این مقاله، ابتدا توضیحاتی در مورد روش‌های استخراج ویژگی پرکاربردی که روش پیشنهادی خود را با آن‌ها مورد مقایسه قرار داده‌ایم، بیان می‌شود. سپس روش پیشنهادی معرفی خواهد شد. نتایج آزمایش‌ها و تحلیل آن‌ها در ادامه بیان شده اند و در نهایت، نتایج به دست آمده بیان گردیده‌اند.

۲- روش‌های استخراج ویژگی نظارت شده

ما روش پیشنهادی خود در این مقاله را با سه گروه از روش‌های استخراج ویژگی مقایسه کرده‌ایم. گروه اول، شامل روش‌هایی است که از میانگین دسته‌ها و یا میانگین وزن‌دار نمونه‌ها برای تولید ماتریس‌های پراکندگی درون دسته‌ای و بین دسته‌ای استفاده می‌کنند و در حقیقت اساس کار آن‌ها بر مبنای جدایی‌پذیری دسته‌ها است. گروه دوم، روش‌هایی هستند که هدف اصلی آن‌ها حفظ ساختار محلی داده در فضای کاهش

پردازش تصاویر از مسائل پر کاربرد مطرح شده در سال‌های اخیر می‌باشد [۱]. در سال‌های اخیر، تصاویر ابرطیفی با فراهم نمودن حجم زیاد اطلاعات طیفی، امکان تفکیک بین عناصر مختلف را با جزئیات دقیق فراهم آورده‌اند. اگر تعداد نمونه‌های آموزشی را ثابت در نظر بگیریم، با افزایش ابعاد داده، صحت طبقه‌بندی ابتدا افزایش و سپس کاهش می‌یابد. این مسئله به پدیده هیوز معروف بوده و یکی از چالش‌های اساسی در تحلیل و طبقه‌بندی تصاویر ابرطیفی است [۲]. کاهش تعداد باندهای طیفی (ویژگی‌ها) علاوه بر تخفیف پدیده‌ی هیوز باعث کاهش پیچیدگی محاسباتی نیز خواهد شد. به طور کلی، روش‌های کاهش ویژگی در دو گروه جای می‌گیرند: انتخاب ویژگی و استخراج ویژگی. هدف فرآیند انتخاب ویژگی، انتخاب موثرترین زیرمجموعه از ویژگی‌ها بر طبق یک تابع معیار و یک الگوریتم جستجو است [۳-۴]. موضوع مورد بررسی ما در این مقاله، استخراج ویژگی است. استخراج ویژگی می‌تواند با استفاده از نمونه‌های آموزشی و به صورت نظارت شده [۵] و یا به صورت بدون نظارت (بدون استفاده از نمونه‌های آموزشی) انجام گیرد. تعدادی از روش‌های بدون نظارت در [۸-۶] پیشنهاد شده‌اند. گروه اول اشاره شده (روش‌های با نظارت) به دلیل در نظر گرفتن جدایی‌پذیری دسته‌ها، مناسب برای کاربردهای طبقه‌بندی‌اند. اما گروه دوم (روش‌های بدون نظارت) تفکیک‌پذیری بین دسته‌ها را در نظر نگرفته و بر اساس معیارهای دیگری از جمله حداقل کردن مربع خطای بازسازی سیگنال عمل می‌کنند و در نتیجه برای کاربردهای طبقه‌بندی ممکن است مفید واقع نشوند. مرور نسبتاً کاملی از روش‌های مختلف انتخاب و استخراج ویژگی در [۹] بیان شده است.

روش استخراج ویژگی پیشنهاد شده در این مقاله، نظارت شده بوده و دارای کارایی خوبی در تعداد نمونه‌های آموزشی محدود است. ما این روش را جاسازی خط ویژگی وزن‌دار^(۲) (WFLE) می‌نامیم. روش پیشنهادی WFLE با استفاده از مفاهیم خط ویژگی^(۳) (FL)، تولید نمونه‌ی آموزشی مجازی می‌کند و از این نمونه‌ها به شکل وزن‌دار برای تولید ماتریس‌های پراکندگی درون دسته‌ای و بین دسته‌ای بهره می‌برد. ایده اساسی روش پیشنهادی در نحوه‌ی وزن‌دهی به این نمونه‌های آموزشی است. وزنی که به هر نمونه‌ی آموزشی مجازی تعلق می‌گیرد به ماهیت نمونه‌های آموزشی تشکیل دهنده‌ی آن بستگی دارد. اگر نمونه‌های آموزشی دارای وضعیت غیرعادی و نامطلوبی باشند، در طی فرآیند استخراج ویژگی، باید اصلاح بیشتری بر روی آن‌ها

$$J(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{S}_b^\Phi \mathbf{w}}{\mathbf{w}^T \mathbf{S}_w^\Phi \mathbf{w}} \quad (4)$$

که \mathbf{S}_b^Φ و \mathbf{S}_w^Φ به ترتیب ماتریس پراکندگی بین-دسته‌ای و درون-دسته‌ای در فضای ویژگی نگاشت یافته هستند. روش GDA هم مانند LDA تنها قادر به استخراج $1 - c$ ویژگی بوده ولی دارای کارایی بهتری نسبت به آن است.

استخراج ویژگی وزن دار غیرپارامتریک (NWFE) [۱۲]: ایده‌ی اساسی روش NWFE، استفاده از میانگین وزن دار و تعریف ماتریس‌های پراکندگی درون-دسته‌ای و بین-دسته‌ای غیرپارامتریک جدید برای استخراج بیش از $1 - c$ ویژگی است. ماتریس‌های پراکندگی در NWFE به شکل زیر تعریف شده‌اند:

$$\mathbf{S}_b = \sum_{i=1}^c p_i \sum_{j=1}^c \sum_{l=1}^{n_i} \frac{\lambda_l^{(i,j)}}{n_i} \left(x_l^{(i)} - M_j(x_l^{(i)}) \right) \left(x_l^{(i)} - M_j(x_l^{(i)}) \right)^T \quad (5)$$

$$\mathbf{S}_w = \sum_{i=1}^c p_i \sum_{l=1}^{n_i} \frac{\lambda_l^{(i,i)}}{n_i} \left(x_l^{(i)} - M_i(x_l^{(i)}) \right) \left(x_l^{(i)} - M_i(x_l^{(i)}) \right)^T \quad (6)$$

در روابط فوق، $x_l^{(i)}$ نمونه l ام از دسته‌ی i ام است. همچنین، $\lambda_l^{(i,j)}$ وزن ماتریس پراکندگی و $M_j(x_l^{(i)})$ میانگین وزن دار $x_l^{(i)}$ در دسته j است که به صورت زیر تعریف می‌گردند:

$$\lambda_l^{(i,j)} = \frac{\text{dist}(x_l^{(i)}, M_j(x_l^{(i)}))^{-1}}{\sum_{t=1}^{n_i} \text{dist}(x_t^{(i)}, M_j(x_t^{(i)}))^{-1}} \quad (7)$$

$$M_j(x_l^{(i)}) = \sum_{k=1}^{n_i} w_{lk}^{(i,j)} x_k^{(j)}, w_{lk}^{(i,j)} = \frac{\text{dist}(x_l^{(i)}, x_k^{(j)})^{-1}}{\sum_{t=1}^{n_j} \text{dist}(x_t^{(i)}, x_t^{(j)})^{-1}} \quad (8)$$

که $\text{dist}(a, b)$ فاصله اقلیدسی بین a و b است. روش NWFE با محاسبه‌ی میانگین وزن دار و تولید ماتریس‌های پراکندگی غیرپارامتریک، باعث بهبود نسبت به روش‌های قبلی می‌شود.

۲-۲-۲ گروه دوم: NPE, LPP

NPE و LPP دو روش معروف و موثر از مجموعه روش‌های یادگیری خمینه^۱ هستند که برای حفظ ساختار محلی داده استفاده می‌شوند. هر کدام از این دو روش می‌توانند به شکل

یافته است. گروه سوم، روش‌هایی هستند که از نمونه‌های آموزشی مجازی تولید شده با مفهوم FL بهره می‌گیرند.

۲-۲-۱ گروه اول: LDA, GDA, NWFE

هر سه روش گروه اول، از معیار جدایی‌پذیری دسته‌ها برای استخراج ویژگی استفاده می‌کنند. این روش‌ها، نسبت ماتریس پراکندگی بین-دسته‌ای (\mathbf{S}_b) را به ماتریس پراکندگی درون-دسته‌ای (\mathbf{S}_w) بیشینه می‌کنند. ویژگی‌های استخراج شده در این روش‌ها با حداکثر کردن معیار $\text{tr}(\mathbf{S}_w^{-1} \mathbf{S}_b)$ به دست می‌آیند.

تحلیل تمییز خطی (LDA) [۱۰]: فرض که d و n به ترتیب بعد داده‌ی اولیه و بعد فضای تبدیل یافته باشند، داده بعد بالا و داده بعد پایین به ترتیب با $\{x_i\}_{i=1}^N$ و $\{z_i\}_{i=1}^N$ نمایش داده شوند و N تعداد کل نمونه‌های داده باشد. نگاشت داده از $x_i \in \mathbb{R}^d$ به \mathbb{R}^d با یک تبدیل خطی $z_i = W^T x_i$ انجام می‌شود. برای به دست آوردن w که بردار تشکیل دهنده‌ی ماتریس تبدیل $\mathbf{W}_{d \times n} = (w_1, w_2, \dots, w_n)$ است، باید مسئله‌ی بهینه‌سازی زیر حل شود [۱۰]:

$$w = \arg \max_w \frac{w^T \mathbf{S}_b w}{w^T \mathbf{S}_w w} \quad (1)$$

در LDA، ماتریس‌های پراکندگی درون-دسته‌ای و بین-دسته‌ای به شکل زیر تعریف می‌شوند:

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

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

در رابطه‌ی فوق، تعداد دسته‌ها برابر c است. n_k ، \mathbf{m}_k و \mathbf{m} و $x_{i,k}$ به ترتیب تعداد نمونه‌های آموزشی دسته k ام، میانگین دسته k ام، میانگین کل نمونه‌های آموزشی و نمونه‌ی i ام در دسته k ام هستند. روش LDA حداکثر قادر به استخراج $1 - c$ ویژگی است. ماتریس پراکندگی درون-دسته‌ای، با وجود نمونه‌های آموزشی محدود، منفرد شده و در این حالت، LDA کارایی ضعیفی خواهد داشت.

تحلیل تمییز تعمیم یافته (GDA) [۱۱]: از طریق نگاشت داده از فضای ورودی ($x \in \mathbb{R}^d$) به فضای ویژگی با بعد بالاتر ($\Phi(x) \in \mathbb{H}$)، روش خطی LDA به روش غیرخطی GDA بسط داده می‌شود. برای به دست آوردن ماتریس تبدیل برای استخراج n ($n < d$) ویژگی، باید تابع هزینه‌ی زیر را بیشینه نمود:

$$w_{ij} = \begin{cases} 1 & x_i, x_j \text{ belong to a class} \\ 0 & x_i, x_j \text{ do not belong to a class} \end{cases}$$

در رابطه‌ی فوق، $n = \sum_{c=1}^{n_c} n_{tc}$ تعداد کل نمونه‌های آموزشی، n_c تعداد دسته‌ها و n_{tc} تعداد نمونه‌های آموزشی دسته c ام است. اگر برای داده‌ی با بعد بالا در فضای اولیه و داده‌ی با بعد کاهش یافته در فضای ثانویه، به ترتیب فرم‌های ماتریسی \mathbf{X} و \mathbf{Y} را به شکل زیر در نظر بگیریم:

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]_{d \times n} \quad (10)$$

$$\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n]_{d \times n} \quad (11)$$

می‌توان تابع Φ را به شکل ماتریسی بازنویسی نمود:

$$\Phi = \text{tr}(\mathbf{Y}\mathbf{G}\mathbf{Y}^T) \\ \min_{\mathbf{A}} [\Phi = \text{tr}(\mathbf{A}\mathbf{X}\mathbf{G}\mathbf{X}^T \mathbf{A}^T)] \quad (12)$$

که در روابط فوق، $\mathbf{Y} = \mathbf{A}\mathbf{X}$ و $\mathbf{G} = \mathbf{D} - \mathbf{W}$ است. \mathbf{W} ماتریس شباهت است که عناصر آن را $w_{ij} (i = 1:n, j = 1:n)$ تشکیل می‌دهند و \mathbf{D} یک ماتریس قطری است که المان‌های روی قطر آن، مجموع سطری ماتریس \mathbf{W} می‌باشند. برای استخراج m ویژگی از d ویژگی اولیه ($m < d$)، تنها کافی است که m بردار ویژه متناظر با m تا کوچک‌ترین مقادیر ویژه ماتریس $\mathbf{X}\mathbf{G}\mathbf{X}^T$ را داخل ماتریس \mathbf{A} بریزیم. در آن صورت، هر نمونه در فضای کاهش یافته، یک بردار m بعدی به شکل $\mathbf{y}_{m \times 1} = \mathbf{A}_{m \times d} \mathbf{x}_{d \times 1}$ خواهد بود.

روش NPE نظارت شده [۱۳]: در این روش، هر نمونه از ترکیب خطی n تا از نزدیک‌ترین همسایه‌های هم دسته‌اش ساخته می‌شود. هر نمونه (\mathbf{x}_i) در فضای اولیه با بعد بالا، می‌تواند از ترکیب خطی نمونه‌های هم دسته‌اش ساخته شود. البته، هر چه نمونه‌ی هم دسته‌اش به آن نزدیک‌تر باشد، وزن بیشتری (مشارکت بیشتری) در ساخت \mathbf{x}_i به آن تعلق می‌گیرد. خطای بازسازی \mathbf{x}_i از همسایه‌هایش، باید حداقل شود. به عبارت دیگر خواهیم داشت:

$$\min \sum_{i=1}^n \|\mathbf{x}_i - \sum_{j=1}^n b_{ij} \mathbf{x}_j\|^2 \quad (13)$$

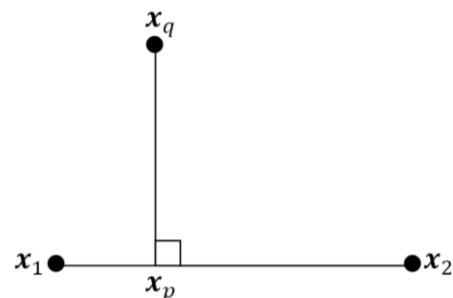
که

بدون نظارت و یا نظارت شده پیاده‌سازی شوند. روش LPP ساختار محلی داده را با استفاده از یک گراف مجاورت^{۱۱} حفظ می‌کند، در حالی که NPE خطای بازسازی را حداقل می‌کند. در فرم‌های نظارت شده NPE و LPP، تنها نمونه‌های هم دسته برای ساخت گراف در نظر گرفته می‌شوند. به عبارت دیگر، تنها شباهت بین نمونه‌های هم دسته محاسبه و فاصله‌ی بین نمونه‌های غیر هم دسته، ارزیابی نمی‌شود. وقتی دو نمونه \mathbf{x}_i و \mathbf{x}_j به یک دسته تعلق داشته باشند، مقادیر در ماتریس شباهت^{۱۲} برابر یک یا یک تابع گاوسی $\exp\left(\frac{-\|\mathbf{x}_i - \mathbf{x}_j\|^2}{t}\right)$ در نظر گرفته می‌شود و در غیر این صورت، مقداری برابر صفر به آن نسبت داده می‌شود. از آنجایی که هدف ما در این تحقیق، بررسی روش‌های استخراج ویژگی نظارت شده است، NPE و LPP نظارت شده را بررسی و در نمونه‌ی آموزشی محدود مورد مقایسه قرار خواهیم داد.

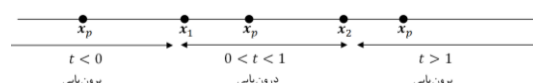
روش LPP نظارت شده [۱۴]: با فرض این که داده در فضای ورودی را \mathbf{x}_i (یک بردار d بعدی) و $\mathbf{y}_i = \mathbf{A}\mathbf{x}_i$ را تبدیل یافته‌ی \mathbf{x}_i در فضای کاهش یافته در نظر بگیریم، ماتریس تبدیل در روش LPP از حل مسئله بهینه‌سازی زیر به دست می‌آید:

$$\min (\Phi = \sum_{i=1}^n \sum_{j=1}^n \|\mathbf{y}_i - \mathbf{y}_j\|^2 w_{ij}) \quad (9)$$

که



شکل ۱: نمایش تصویر نقطه مفروض \mathbf{x}_q بر \mathbf{x}_p



شکل ۲: حالت‌های مختلف درون‌یابی و برون‌یابی در NFL

$$b_{ij} = \begin{cases} \frac{d^{-1}(x_i, x_j)}{\sum_{k=1}^{n_{tc}} d^{-1}(x_i, x_k)} & x_i, x_j \text{ belong to class } c \text{ and } i \neq j \\ \cdot & x_i, x_j \text{ do not belong to a class} \end{cases}$$

در رابطه‌ی فوق، $d(x_i, x_j) = (x_i - x_j)^T(x_i - x_j)$ است. وزن‌های b_{ij} نرمالیزه شده‌اند تا شرط $\sum_{j=1}^n b_{ij} = 1$ فراهم شود. همان‌طور که در فضای اولیه، هر نمونه از ترکیب خطی نمونه‌های هم دسته‌اش ساخته می‌شود (به صورتی که نمونه‌های همسایه‌ی نزدیک‌تر، نقش بیشتری در این ترکیب داشته باشند)، در فضای کاهش یافته هم، هر نمونه می‌تواند از نمونه‌های هم دسته‌اش با همان ضرایب وزنی، بازسازی شود. به این ترتیب، ساختار محلی داده در فضای کاهش یافته، حفظ خواهد شد. در این صورت خواهیم داشت:

$$\psi = \sum_{i=1}^n \|y_i - \sum_{j=1}^n b_{ij} y_j\|^2 \quad (14)$$

که به شکل ماتریس زیر باز نویسی خواهد شد:

$$\psi = \|Y(I - B)\|_F^2 \quad (15)$$

نرم فروبنیوس است. با ساده‌سازی عبارت فوق خواهیم داشت:

$$\psi = \text{tr}[Y(I - B)(I - B)^T Y^T] \quad (16)$$

با در نظر گرفتن $Y = AX$ و $M = (I - B)(I - B)^T$ داریم:

$$\min_A [\psi = \text{tr}(AXMX^T A^T)] \quad (17)$$

برای استخراج m ویژگی از d ویژگی اولیه ($m < d$)، ماتریس تبدیل A از m بردار ویژه متناظر با m تا از کوچک‌ترین مقادیر ویژه ماتریس $AXMX^T$ ساخته می‌شود.

روش NPE خطای بازسازی را حداقل می‌کند، در حالی که LPP فاصله‌ی بین نقاط همسایه را حداقل می‌کند. لازم به ذکر است که روش‌های NPE و LPP که هدف اصلی آن‌ها، حفظ ساختار محلی داده است، برای کاهش ویژگی و بازسازی داده خوب کار می‌کنند و ممکن است برای کاربردهای طبقه‌بندی، به اندازه کافی مطلوب نباشند.

۲-۲-۳- گروه سوم: NFLE

در این قسمت، ابتدا به توصیف مفاهیم خط ویژگی و استفاده از آن در طبقه‌بندی می‌پردازیم. سپس بیان می‌کنیم که چگونه از این مفاهیم، می‌توان برای استخراج ویژگی بهره برد.

نزدیک‌ترین خط ویژگی (NFL^{۱۳}): روشی است برای طبقه‌بندی که در حقیقت بهبود یافته‌ی طبقه‌بند نزدیک‌ترین همسایگی (NN) است و اولین بار به عنوان یک طبقه‌بند مناسب برای مسائل نمونه‌ی آموزشی محدود در سال ۱۹۹۹ ارائه شد [۱۶]. اجازه دهید x_q یک نمونه مفروض و x_1 و x_2 دو نمونه‌ی آموزشی هم دسته باشند. خط واصل $\overline{x_1 x_2}$ ، خط ویژگی (FL) نامیده می‌شود. از نمونه x_q بر FL تشکیل شده از نقاط x_1 و x_2 خط عمودی رسم می‌کنیم و پای عمود را x_p می‌نامیم. در این حالت، $\|x_q - x_p\|$ فاصله‌ی خط ویژگی^{۱۴} نامیده می‌شود. مقدار بردار x_p با استفاده از بردارهای x_1 ، x_2 و x_q قابل محاسبه است:

$$x_p = x_1 + t(x_2 - x_1) \quad (18)$$

t که پارامتر موقعیت^{۱۵} نامیده می‌شود، با توجه به رابطه‌ی تعامد $x_1 x_2 \perp x_p x_q$ قابل محاسبه خواهد بود:

$$\begin{aligned} x_1 x_2 \perp x_p x_q &\rightarrow (x_p - x_q) \cdot (x_2 - x_1) = 0 \\ \rightarrow t &= \frac{(x_q - x_1)^T (x_2 - x_1)}{(x_2 - x_1)^T (x_2 - x_1)} \end{aligned} \quad (19)$$

شکل ۱ موقعیت نمونه‌ها را نشان می‌دهد. خط واصل بین x_1 و x_2 ، $\overline{x_1 x_2}$ ، نه تنها تغییرات احتمالی x_1 و x_2 را فراهم می‌کند، بلکه نمونه‌های آموزشی مجازی را تولید می‌کند. با توجه به مقدار پارامتر t ، درون‌یابی یا برون‌یابی انجام می‌پذیرد. به عبارت دیگر خواهیم داشت:

$$\begin{aligned} x_p &= x_1 + t(x_2 - x_1) \\ t = 1 &\rightarrow x_p = x_2 \\ t = 0 &\rightarrow x_p = x_1 \\ 0 < t < 1 &\rightarrow x_p = \text{linear interpolation } (x_1, x_2) \\ t < 0 \text{ or } t > 1 &\rightarrow x_p \\ &= \text{linear extrapolation } (x_1, x_2) \end{aligned}$$

مطابق روابط فوق، زمانی که $0 < t < 1$ ، x_p یک درون‌یابی خطی بین x_1 و x_2 و زمانی که $t < 0$ یا $t > 1$ ، x_p یک برون‌یابی خطی بین x_1 و x_2 است. این موضوع در شکل ۲ به خوبی نشان داده شده است.

به جای عبور یک خط از بین دو نقطه برای تولید FL، می‌توان برای شرح دقیق‌تر تغییرات بین نقاط، یک منحنی با

درجه‌ی بالاتر از میان نقاط عبور داد که البته پیچیدگی در این حالت بیشتر خواهد شد.

در بسیاری از مقالات از NFL به عنوان طبقه‌بند استفاده کرده‌اند [۱۶-۱۷]، اما مقالات محدودی هم در سال‌های اخیر ارائه شده‌اند که از مفاهیم NFL در فاز استخراج ویژگی استفاده کرده‌اند.

جاسازی نزدیک‌ترین خط ویژگی (NFLE) [۱۵]: در جدیدترین روش پیشنهادی که در سال ۲۰۱۴ ارائه شده است، ابتدا در صورتی که اندازه مجموعه نمونه‌ی آموزشی کوچک باشد، با استفاده از تبدیل PCA^{۱۶} یک کاهش بعد اولیه بر روی داده انجام می‌شود. سپس از تصویر نمونه‌های آموزشی بر روی FLها برای تولید ماتریس‌های پراکندگی S_W و S_B استفاده می‌گردد. در این روش که NFLE نامیده شده، به جای این‌که فاصله‌ی نمونه‌ها تا میانگین دسته‌ها یا میانگین‌های کلی و محلی محاسبه شود، فاصله‌ی آن نمونه تا تصویر آن، بر روی FLهای تولید شده توسط مجموعه نمونه‌های آموزشی، محاسبه می‌گردد. اگر در دسته C ، n_{tc} نمونه‌ی آموزشی موجود باشد، فاصله‌ی هر نمونه از دسته C تا $\frac{(n_{tc}-1) \times (n_{tc}-2)}{2} = \binom{n_{tc}-1}{2}$ تصویر آن از جفت FLهای ساخته شده، محاسبه می‌شود. به این ترتیب، ماتریس‌های S_W و S_B به شکل زیر قابل محاسبه خواهند بود [۱۵]:

$$S_W = \sum_{i=1}^n \sum_{j \in P_i} (\mathbf{x}_i - \mathbf{x}_{ij})(\mathbf{x}_i - \mathbf{x}_{ij})^T \quad (20)$$

$$S_B = \sum_{i=1}^n \sum_{k \in R_i} (\mathbf{x}_i - \mathbf{x}_{ik})(\mathbf{x}_i - \mathbf{x}_{ik})^T \quad (21)$$

که n تعداد کل نمونه‌های آموزشی است. \mathbf{x}_{ij} تصویر \mathbf{x}_i بر روی FLهای است که از نمونه‌های هم دسته با \mathbf{x}_i شکل گرفته‌اند و \mathbf{x}_{ik} تصویر \mathbf{x}_i بر روی FLهایی است که از نمونه‌هایی که دارای برچسب متفاوت با \mathbf{x}_i می‌باشند، شکل گرفته‌اند. اگر L_{x_i} برچسب نمونه‌ی \mathbf{x}_i را نشان دهد، مجموعه‌های P_i و R_i به شکل زیر تعریف می‌شوند:

$$P_i = \{j \mid L_{x_i} = L_{x_{ij}}\} \quad (22)$$

$$R_i = \{k \mid L_{x_i} \neq L_{x_{ik}}\} \quad (23)$$

به عبارت دیگر، P_i مجموعه اندیس‌های نقاط تصویر \mathbf{x}_i بر FLهای هم دسته با \mathbf{x}_i و R_i مجموعه اندیس‌های نقاط تصویر \mathbf{x}_i بر FLهای غیر هم دسته با \mathbf{x}_i است. در دو مقاله‌ای که در سال‌های قبل‌تر از این، ارائه شده نیز به نحوی از مفاهیم NFL

استفاده گردیده است. در [۱۸] که در سال ۲۰۰۷ معرفی شد، تنها پراکندگی‌های درون دسته‌ای، درون متریک NFL جاساز شده است. در [۱۹] که در سال ۲۰۱۰ ارائه شد، از هر دو اطلاعات درون دسته‌ای و بین دسته‌ای استفاده شده، اما به جای تولید ماتریس‌های پراکندگی، میانگین مربعات فاصله‌ی بین نمونه‌های آموزشی و نقاط تصویر نظیرشان بر روی FLها در فضای تبدیل یافته حداقل گردیده است.

۳- روش پیشنهادی (WFLE)

روش پیشنهادی ما در این تحقیق، تولید ماتریس‌های پراکندگی درون دسته‌ای و بین دسته‌ای با استفاده از مفاهیم NFL اما به شکل وزن‌دار است و WFLE نامیده می‌شود. در WFLE، با توجه به ماهیت نمونه‌های آموزشی، وزن مناسب به تصویر هر نمونه در تولید ماتریس‌های S_W و S_B داده می‌شود. توضیحات بیشتر در مورد جزئیات این روش در ادامه بیان شده است. نمونه‌های آموزشی در یکی از چهار گروه زیر جای می‌گیرند:

۱- نمونه‌های هم دسته شبیه

۲- نمونه‌های غیر هم دسته متفاوت

۳- نمونه‌های هم دسته متفاوت

۴- نمونه‌های غیر هم دسته شبیه

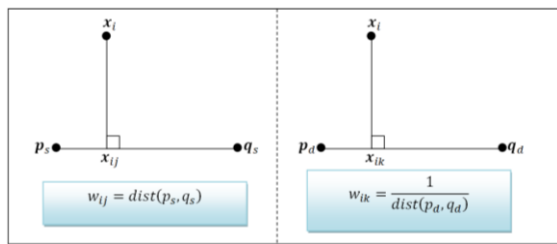
منظور از نمونه‌های شبیه، نمونه‌هایی است که فاصله‌ی کمی بین آن‌ها وجود دارد و نمونه‌های متفاوت، نمونه‌هایی هستند که فاصله‌ی زیادی بین آن‌ها است. فاصله‌ی مورد استفاده در این‌جا، فاصله‌ی اقلیدسی است. دو دسته اول بیان شده، وضعیت خوب و طبیعی دارند که مطلوب است. اما دو دسته بعدی، دارای وضعیت غیرعادی هستند که مطلوب ما نیست و در فرآیند طبقه‌بندی، منجر به خطا در برچسب زدن به نمونه‌های آزمایشی خواهند شد. نمونه‌هایی از این ۴ گروه داده آموزشی برای حالت دو بعدی و دو دسته، در شکل ۳ نشان داده شده است. ما وزن‌ها را در روش پیشنهادی WFLE به گونه‌ای به دست می‌آوریم که دو دسته نامطلوب اشاره شده، در فضای تبدیل شده (کاهش یافته)، بیشتر اصلاح شوند. ماتریس‌های پراکندگی درون دسته‌ای (S_W) و بین دسته‌ای (S_B) در WFLE مطابق روابط زیر تعریف می‌شوند:

$$S_W = \sum_{i=1}^n \sum_{j \in P_i} w_{ij} (\mathbf{x}_i - \mathbf{x}_{ij})(\mathbf{x}_i - \mathbf{x}_{ij})^T \quad (24)$$

روش NFLE، در ابتدای امر یک کاهش ویژگی اولیه با استفاده از PCA انجام می‌دهد. اما از آنجایی که تبدیل PCA، ویژگی‌ها را بر مبنای قدرت (واریانس) آن‌ها مرتب می‌کند که این وضعیت ایجاد شده در فضای ویژگی ممکن است برای کاربردهای طبقه‌بندی چندان مفید نباشد، ما این مرحله کاهش ویژگی اولیه را در روش پیشنهادی خود، WFLE، حذف نموده‌ایم. ما پس از تولید ماتریس‌های پراکندگی S_b و S_w برای جلوگیری از منفرد شدن ماتریس پراکندگی درون دسته‌ای، آن را به شکل زیر تنظیم^{۱۷} می‌کنیم:

$$S_w = 0.5S_w + 0.5diag(S_w) \quad (28)$$

سپس با حداکثر کردن معیار فیشر، $tr(S_w^{-1}S_b)$ ، ماتریس تبدیل WFLE را به دست خواهیم آورد. در حقیقت، اساس وزن‌دهی در روش پیشنهادی WFLE بر مبنای زیر شکل گرفته است:

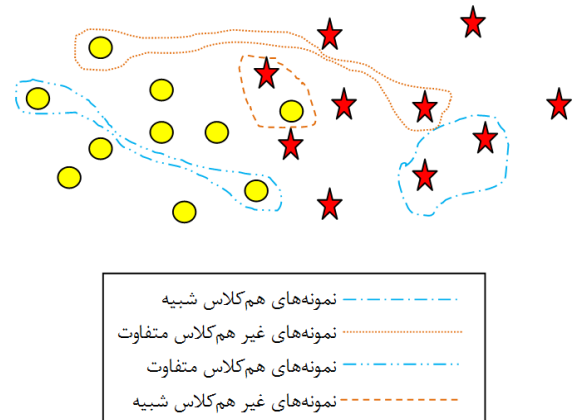


شکل ۴: تصویر x_i بر روی FL‌های هم دسته و غیر هم دسته

اگر نمونه‌های آموزشی در وضعیت مطلوبی می‌باشند (نمونه‌های هم دسته شبیه یا نمونه‌های غیر هم دسته متفاوت)، برای طبقه‌بندی مناسب بوده و نیاز به اصلاح چندان در فضای تبدیل شده (کاهش یافته) ندارند. اما اگر نمونه‌های آموزشی در وضعیت نامطلوبی هستند که منجر به ایجاد خطا در فرآیند طبقه‌بندی می‌گردد (نمونه‌های هم دسته متفاوت یا نمونه‌های غیر هم دسته شبیه)، باید در فرآیند استخراج ویژگی، اصلاح بیشتری بر روی آن‌ها انجام شود تا منجر به تولید نمونه‌های مطلوب‌تر در فضای کاهش یافته شود. در نتیجه، به این نمونه‌ها در فرآیند کاهش ویژگی، در تولید ماتریس‌های پراکندگی، وزن بیشتری اختصاص داده خواهد شد. به عبارت دیگر، برای به دست آوردن ماتریس تبدیل استخراج ویژگی، که باید پراکندگی‌های درون دسته‌ای را حداقل و پراکندگی‌های بین دسته‌ای را حداکثر کنیم، با دادن وزن بیشتر به نمونه‌های غیرعادی در تولید ماتریس‌های

$$S_b = \sum_{i=1}^n \sum_{k \in R_i} w_{ik} (x_i - x_{ik})(x_i - x_{ik})^T \quad (25)$$

فرض کنید p_s و q_s دو نمونه‌ی هم دسته با x_i باشند. در این صورت، تصویر نمونه‌ی x_i بر FL حاصل از p_s و q_s است. اگر p_s و q_s متفاوت باشند، یعنی فاصله‌ی اقلیدسی بین آن‌ها زیاد باشد، $dist(p_s, q_s)$ بزرگ باشد، جزء گروه غیرعادی (نامطلوب) نمونه‌ها محسوب می‌شوند و باید وزن بیشتری برای اصلاح آن‌ها اختصاص داده شود. اگر p_s و q_s به هم شبیه باشند، یعنی فاصله‌ی بین آن‌ها کم باشد، $dist(p_s, q_s)$ کوچک باشد، پس باید وزن کمتری به آن‌ها در تولید ماتریس پراکندگی اختصاص داده شود. پس رابطه‌ی مستقیم بین فاصله‌ی نمونه‌ها و ضریب وزنی وجود دارد. بر این اساس، وزن‌های w_{ij} از رابطه‌ی زیر به دست می‌آیند:



شکل ۳: نمونه‌هایی از چهار نوع دسته‌بندی داده

$$w_{ij} = dist(p_s, q_s) = (p_s - q_s)^T (p_s - q_s) \quad (26)$$

اگر p_a و q_a دو نمونه‌ی غیر هم دسته باشند، تصویر x_i بر روی FL حاصل از p_a و q_a نامیده می‌شود. اگر p_a و q_a به هم شبیه باشند، یعنی $dist(p_a, q_a)$ کوچک باشد، جزء گروه غیرعادی و نامطلوب جای می‌گیرند و باید وزن بیشتری برای اصلاح به آن‌ها اختصاص داد. بر عکس، اگر آن‌ها متفاوت باشند، $dist(p_a, q_a)$ بزرگ باشد، وضعیت مطلوب است و وزن کمتری به آن‌ها تعلق می‌گیرد. پس رابطه‌ی معکوس بین فاصله‌ی نمونه‌ها و ضریب وزنی w_{ik} وجود دارد:

$$w_{ik} = \frac{1}{dist(p_a, q_a)} = [(p_a - q_a)^T (p_a - q_a)]^{-1} \quad (27)$$

نقطه x_i و تصویر آن بر روی FL‌های هم دسته و غیر هم دسته آن در شکل ۴ نشان داده شده است.

پراکندگی، آن‌ها را به سمت اصلاح شدن بیشتر سوق می‌دهیم. به این ترتیب، روش پیشنهادی دارای دو ویژگی موثر است:

- ۱- به دلیل استفاده از مفهوم FL، که منجر به تولید نمونه‌ی آموزشی مجازی می‌شود، در شرایط نمونه‌ی آموزشی محدود، دارای کارایی خوبی است.
- ۲- به دلیل به کارگیری روش وزن‌دهی بر اساس ماهیت نمونه‌های آموزشی مطلوب و غیر مطلوب (عادی و غیرعادی)، منجر به بهبود فرآیند جداسازی دسته‌ها و افزایش صحت طبقه‌بندی خواهد شد.

استفاده از روش‌های بر مبنای NFL در تعداد نمونه‌ی آموزشی بالا، به دو دلیل چندان توصیه نمی‌شود: ۱- افزایش پیچیدگی محاسباتی ۲- اگر به تعداد کافی نمونه‌ی آموزشی واقعی در دسترس باشد، دیگر استفاده از نمونه‌های آموزشی مجازی، بی‌معنا و احتمالاً ناکارآمد خواهد بود. بلوک دیگرام روش پیشنهادی در شکل ۵ مشاهده می‌شود.

۴- نتایج آزمایش‌ها

در ابتدا در مورد داده‌های ابرطیفی مورد استفاده در آزمایش‌ها و معیارهای ارزیابی برای مقایسه‌ی روش‌های استخراج ویژگی توضیحاتی داده و سپس به بیان آزمایش‌ها و تحلیل آن‌ها می‌پردازیم.

۴-۱- داده‌ها و معیارهای ارزیابی

در این مقاله، از سه تصویر ابرطیفی برای انجام آزمایش‌ها استفاده شده است: ایندیانا^{۱۸}، دانشگاه پاپویا^{۱۹} و سالیانس-A^{۲۰}. در ادامه به توضیح این داده‌های ابرطیفی می‌پردازیم.

ایندیانا: تصویر ابرطیفی مربوط به منطقه ای کشاورزی- جنگلی است که توسط سنجنده AVIRIS^{۲۱} در ژوئن ۱۹۹۲ از یک سایت آزمایشی در شمال شرقی ایالت ایندیانا گرفته شده است. این تصویر حاوی ۱۶ دسته، ۲۲۰ باند باریک با پهنای ۱۰ نانومتر در بازه فرکانسی ۰٫۴ تا ۲٫۵ میکرومتر، ۱۴۵×۱۴۵ پیکسل و صحت مکانی ۲۰ متر بر پیکسل و صحت رادیومتریک ۸ بیت است. بعد از حذف ۲۰ کانال نویزی، آزمایش‌ها بر روی ۲۰۰ باند باقیمانده انجام شده است. ما در آزمایش‌های خود از میان ۱۶ دسته داده‌ی ایندیانا، از ۱۰ دسته این داده که بیشترین

تعداد نمونه‌ی‌های مرجع را دارا هستند، و در بسیاری از تحقیقات مرسوم هستند، استفاده نموده‌ایم.

دانشگاه پاپویا: این تصویر شهری مربوط به ناحیه محوطه دانشگاه پاپویا در ایتالیاست. این تصویر با استفاده از سنجنده نوری ROSIS^{۲۲} اخذ شده است. تصویر دانشگاه پاپویا، دارای تفکیک مکانی ۱٫۳ متر بر پیکسل و دارای ۱۱۵ باند طیفی در بازه فرکانسی ۰٫۴۳ تا ۰٫۸۶ میکرومتر است. این تصویر شامل ۹ دسته بوده و ابعاد آن ۳۴۰×۶۱۰ است. بعد از حذف کانال‌های نویزی، آزمایش‌ها بر روی ۱۰۳ باند باقیمانده انجام شده است.

سالیانس-A: یک زیرصحنه کوچک از تصویر سالیانس، تصویر سالیانس-A است. این تصویر شامل ۸۶×۸۳ پیکسل و ۶ دسته از داده‌ی سالیانس است. تصویر سالیانس از دهکده سالیانس در کالیفرنیا، جنوبی در سال ۱۹۹۸ توسط سنجنده AVIRIS اخذ شده است. این تصویر دارای اندازه پیکسل ۳٫۷ متر و ۲۲۴ باند طیفی در بازه فرکانسی ۰٫۴ تا ۲٫۵ میکرومتر، با قدرت تفکیک طیفی ۱۰ نانومتر است. بعد از حذف ۲۰ باند جذب آب، ۲۰۴ باند از این تصویر در آزمایش‌ها استفاده می‌شوند. مشخصات داده‌های ابرطیفی مورد استفاده در آزمایش‌ها در جدول ۱ ذکر شده‌اند.

ما برای داشتن رفتار و مقایسه منصفانه با تمام دسته‌ها، تعداد نمونه‌های آموزشی یکسانی در تمام دسته‌ها در نظر گرفته و باقیمانده‌ی نمونه‌های هر دسته را به عنوان نمونه‌ی آزمایشی استفاده می‌کنیم. نمونه‌های آموزشی به طور تصادفی از سراسر صحنه انتخاب شده‌اند. هر آزمایش، ۱۰ مرتبه تکرار شده و متوسط نتایج به دست آمده، گزارش شده است.

موضوع مورد تحقیق ما در این مقاله، استخراج ویژگی با استفاده از تعداد نمونه‌های آموزشی محدود برای کاربردهای طبقه‌بندی است. در نتیجه، ما از دو طبقه‌بند با رفتار کاملاً متفاوت برای ارزیابی روش‌های استخراج ویژگی استفاده نموده‌ایم. طبقه‌بند مورد استفاده اول، طبقه‌بند ماشین بردار پشتیبان (SVM^{۲۳}) است [۲۰-۲۱]. این طبقه‌بند غیرپارامتریک بوده و به تعداد نمونه‌های آموزشی، حساسیت کم‌تری دارد و به طور کلی برای تعداد نمونه‌های آموزشی محدود، بسیار مناسب و خوش‌رفتار است. طبقه‌بند دیگر مورد استفاده طبقه‌بند بیشینه احتمال (ML^{۲۴}) است. طبقه‌بند ML، پارامتریک بوده و به تعداد نمونه‌های آموزشی بسیار حساس است. به گونه‌ای که برای جلوگیری از منفرد شدن ماتریس کواریانس دسته‌ها، تعداد نمونه‌های آموزشی از تعداد ویژگی‌ها، باید حداقل یکی بیشتر باشد.

ما برای استفاده از طبقه‌بند ML، توزیع معمول گاوسی را برای دسته‌ها در نظر گرفته‌ایم. برای استفاده از طبقه‌بند SVM از کتابخانه LIBSVM بهره برده‌ایم [۲۲]. طبقه‌بند SVM مورد استفاده در انجام آزمایش‌ها در حالت چند دسته با استفاده از روش یکی در برابر یکی و با تابع هسته (کرنل) چندجمله‌ای درجه ۳ با مقادیر پیش فرض LIBSVM است.

معیارهای مختلفی برای ارزیابی کارایی طبقه‌بندها می‌تواند مورد استفاده قرار گیرند که از آن جمله می‌توان به صحت متوسط و اعتبار متوسط طبقه‌بندی اشاره کرد. صحت متوسط، میانگین صحت به دست آمده برای تمام دسته‌ها است. اعتبار متوسط نیز برابر میانگین اعتبار به دست آمده برای تمام دسته‌ها است. اعتبار برای هر دسته این‌گونه تعریف شده است: تعداد نمونه‌هایی که درست طبقه‌بندی شده‌اند، تقسیم بر تعداد کل نمونه‌هایی که به آن دسته تعلق گرفته‌اند.

آزمون McNemars برای بیان معناداری تفاوت بین روش‌های طبقه‌بندی از لحاظ آماری بسیار مفید است [۲۳]. این آزمون برای مقایسه بین یک جفت طبقه‌بند مورد استفاده قرار می‌گیرد. پارامتر Z_{12} به صورت زیر تعریف می‌شود:

$$Z_{12} = \frac{f_{12} - f_{21}}{\sqrt{f_{12} + f_{21}}} \quad (29)$$

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 برعکس ($Z_{12} < 0$).

۴-۲- نتایج و تحلیل آزمایش‌ها

ما روش پیشنهادی خود در این قسمت، یعنی WFLE، را با سه گروه از روش‌های استخراج ویژگی نظارت شده مورد مقایسه قرار می‌دهیم. گروه اول، روش‌های استخراج ویژگی‌ای هستند که بر مبنای جدایی‌پذیری دسته‌ها کار می‌کنند. این روش‌ها با تولید ماتریس‌های پراکندگی درون دسته‌ای و بین دسته‌ای سعی در کمینه کردن پراکندگی‌های درون دسته‌ای و بیشینه کردن پراکندگی‌های بین دسته‌ای دارد. از جمله‌ی این روش‌ها می‌توان به LDA [۱۰]، GDA [۱۱] و NWFE [۱۲] اشاره نمود. این گروه از روش‌ها از میانگین دسته‌ها یا میانگین وزن‌دار نمونه‌ها استفاده کرده و تخمینی از ماتریس پراکندگی (کواریانس) بین

دسته‌ای و درون دسته‌ای را محاسبه می‌کنند. به دلیل نیاز این روش‌ها به تخمین بردارهای میانگین و ماتریس‌های کواریانس، این روش‌ها، حساسیت بیشتری به تعداد نمونه‌های آموزشی دارند. به گونه‌ای که با افزایش تعداد نمونه‌های آموزشی، تخمین بهتری از آمارگان‌ها به دست می‌آید که منجر به افزایش صحت طبقه‌بندی خواهد شد.

گروه دوم روش‌های استخراج ویژگی، روش‌هایی هستند که برای حفظ ساختار محلی داده مورد استفاده قرار می‌گیرند. از جمله این روش‌ها می‌توان به NPE [۱۳] و LPP [۱۴] اشاره نمود که می‌توانند به دو صورت نظارت شده و بدون نظارت، پیاده‌سازی شوند که البته فرم نظارت شده‌ی آن‌ها برای ما مد نظر است. از آنجایی که این روش‌ها، نیازی به تخمین آمارگان ندارند، حساسیت آن‌ها به تعداد نمونه‌های آموزشی کمتر است. از طرفی، به دلیل این‌که این روش‌ها بر مبنای حفظ ساختار محلی داده کار می‌کنند، به نظر می‌رسد، بیشتر برای کاربردهای بازسازی داده مناسب باشند تا کاربردهای طبقه‌بندی. اما به هر حال، زمانی که نمونه‌ی آموزشی به تعداد کافی در اختیار نباشد تا امکان تخمین مناسبی از آمارگان دسته‌ها را فراهم کند، استفاده از چنین روش‌هایی می‌تواند کارایی داشته باشد.

و اما گروه سوم روش‌های استخراج ویژگی که در این جا بررسی می‌شود، روش‌هایی هستند که با تولید نمونه‌ی آموزشی مجازی، سعی در جبران کمبود نمونه‌ی آموزشی دارند. از جمله این روش‌ها می‌توان به NFLE [۱۵] اشاره نمود که با استفاده از مفهوم FL، نمونه‌های آموزشی مجازی تولید کرده و از آن‌ها برای تخمین ماتریس‌های پراکندگی درون دسته‌ای و بین دسته‌ای استفاده می‌کند. به عنوان مثال، اگر در هر دسته، ۱۰ نمونه‌ی آموزشی داشته باشیم، با استفاده از مفهوم FL، $(10^{-1}) = 36 = \frac{9 \times 8}{4}$ نمونه‌ی آموزشی مجازی تولید خواهد شد که این خود صحت تخمین ماتریس‌های پراکندگی و در نتیجه صحت طبقه‌بندی داده با بعد کاهش یافته را بسیار بهبود خواهد بخشید.

روش پیشنهادی ما، WFLE، نیز با استفاده از مفاهیم FL، نمونه‌ی آموزشی مجازی تولید کرده و از آن‌ها برای تخمین ماتریس‌های پراکندگی درون دسته‌ای و بین دسته‌ای استفاده می‌کند. اما نوآوری این روش، همان‌طور که قبلاً هم اشاره شد، در وزن‌دهی به نمونه‌ها در هنگام تولید ماتریس‌های پراکندگی است. اساس این وزن دهی بر مبنای اصلاح بیشتر نمونه‌های نامطلوبی است که وضعیت غیرعادی داشته و منجر به ایجاد خطا در فرآیند طبقه‌بندی می‌گردند. تفاوت دیگر روش WFLE با

NFLE در این است که NFLE، در ابتدای امر یک کاهش ویژگی اولیه با استفاده از PCA انجام می‌دهد و سپس استخراج ویژگی مورد نظر را با استفاده از ماتریس‌های پراکندگی تخمین زده شده از نمونه‌های آموزشی مجازی انجام می‌دهد. این در حالی است که می‌دانیم تبدیل PCA ویژگی‌ها را بر مبنای قدرت (واریانس) آن‌ها مرتب می‌کند و ویژگی‌های با قدرت بالا را حفظ و مابقی را دور می‌ریزد. از آن جایی که ممکن است در میان ویژگی‌های با قدرت کم نیز اطلاعات با ارزشی در مورد جدایی‌پذیری دسته‌ها وجود داشته باشد، که در تبدیل PCA نادیده گرفته شده و این اطلاعات دور ریخته می‌شود، ما در روش پیشنهادی WFLE این مرحله کاهش ویژگی اولیه را روی داده نداریم و یک سره روش استخراج ویژگی پیشنهادی را بر روی داده با بعد کامل اعمال می‌کنیم. ما برای جلوگیری از منفرد شدن ماتریس پراکندگی درون دسته‌ای، همان‌طور که قبلاً نیز بیان شد، آن را تنظیم می‌کنیم.

روش پیشنهادی WFLE با سه گروه از روش‌های استخراج ویژگی بیان شده در بالا، یعنی با روش‌های LDA، GDA، NPE، LPP، NWFE و NFLE مقایسه شده است. ما برای آزمایش‌های خود، از ۱۵ نمونه‌ی آموزشی و طبقه‌بندی‌های SVM و ML استفاده کرده‌ایم. به علاوه با استفاده از ۵ نمونه‌ی آموزشی و طبقه‌بندی SVM نیز آزمایش‌ها انجام گرفته است. لازم به ذکر است که امکان استفاده از طبقه‌بند ML با ۵ نمونه‌ی آموزشی وجود ندارد. چراکه در طبقه‌بند ML برای جلوگیری از منفرد شدن ماتریس کواریانس دسته‌ها، تعداد نمونه‌های آموزشی باید حداقل از تعداد ویژگی‌ها، یکی بیشتر باشد و با ۵ نمونه آموزشی، حداکثر قادر به استفاده از ۴ ویژگی در طبقه‌بند ML خواهیم بود.

خلاصه نتایج به دست آمده از آزمایش‌ها، در جدول ۲ بیان است. صحت متوسط طبقه‌بندی در برابر تعداد ویژگی‌های استخراج شده برای داده‌ی ایندیانا در شکل ۶ نمایش داده شده است. به عنوان نمونه، بالاترین صحت متوسط طبقه‌بندی به ازای ۵ نمونه‌ی آموزشی و طبقه‌بند SVM برای سه داده‌ی ایندیانا، دانشگاه پاپویا و سالیانس-A به ترتیب به ازای ۴، ۵ و ۳ ویژگی استخراجی به دست آمده است. بالاترین صحت و اعتبار به دست آمده دسته‌ها برای هر سه داده‌ی ابرطیفی ایندیانا، دانشگاه پاپویا و سالیانس-A به ترتیب در جدول‌های ۳، ۴ و ۵ بیان شده‌اند. هم‌چنین، نقشه‌ی طبقه‌بندی به دست آمده نظیر این جدول‌ها نیز، به ترتیب در شکل‌های ۷، ۸ و ۹ نشان داده شده‌اند. مقادیر پارامتر Z حاصل از آزمون MCNemars نیز به ازای بالاترین

صحت متوسط به دست آمده برای همه‌ی حالت‌ها و هر سه داده ابرطیفی در جدول ۶ بیان گردیده‌اند. با توجه به نتایج به دست آمده از آزمایش‌ها، می‌توان به نکات زیر اشاره کرد:

۱- برای داده‌ی ایندیانا، در هر ۳ حالت، ۵ نمونه و طبقه‌بند SVM، ۱۵ نمونه و طبقه‌بند SVM، ۱۵ نمونه و طبقه‌بند ML، روش WFLE به سایر روش‌های استخراج ویژگی برتری دارد. از آن جایی که داده‌ی ایندیانا، داده سختی برای طبقه‌بندی است، تفاوت بین روش‌های مناسب و نامناسب، بیشتر در این داده آشکار می‌شود.

۲- برای داده‌ی سالیانس-A، به ازای ۵ نمونه آموزشی، WFLE به سایر روش‌ها برتری دارد. از نظر صحت متوسط طبقه‌بندی، با ۱۵ نمونه و طبقه‌بند SVM، WFLE با NFLE معادل و با ۱۵ نمونه و طبقه‌بند ML، WFLE با همه‌ی روش‌ها جز LDA معادل است. اما با توجه به این‌که با استفاده از ۱۵ نمونه‌ی آموزشی، صحت روش‌ها، معادل هم شده است، برای فهم تفاوت بین این روش‌ها و این‌که کدام یک از نظر آماری بر سایرین برتری دارد از معیار MCNemars استفاده می‌کنیم. اگر روش WFLE را با شماره ۱ و سایر روش‌های NFLE، LDA، NWFE، GDA، NPE و LPP را به ترتیب با شماره‌های ۲ و ۳ و ۴ و ۵ و ۶ و ۷ در نظر بگیریم، با استفاده از این معیار، و محاسبه پارامتر Z برای همه جفت روش‌ها، متوجه می‌شویم که روش WFLE بر سایر روش‌ها برتری دارد یعنی همه‌جا $Z_{ij} > 0$ ($i = 2, \dots, 7$) و در بیشتر موارد $|Z_{ij}| > 1.96$ است که نشان می‌دهد این تفاوت از نظر آماری قابل توجه است.

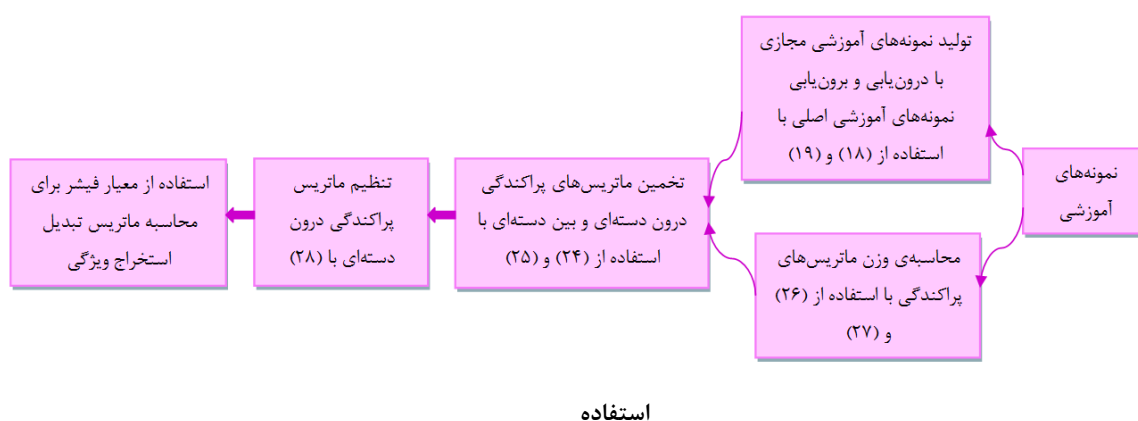
۳- برای داده‌ی پاپویا، با ۱۵ نمونه و طبقه‌بند ML، روش WFLE با NFLE و GDA از نظر صحت متوسط طبقه‌بندی معادل است ولی، با استفاده از معیار MCNemars، WFLE نسبت به همه روش‌ها برتری دارد. با ۵ نمونه و طبقه‌بند SVM، WFLE با GDA، LPP و NPE از نظر صحت طبقه‌بندی معادل است اما با معیار MCNemars، روش WFLE بعد از روش GDA نسبت به سایر روش‌ها برتری دارد. البته مقدار پارامتر Z بین WFLE و GDA برابر $Z_{15} = -1.47$ است، به عبارت دیگر، $|Z_{15}| < 1.96$. هم‌چنین، با ۱۵ نمونه و طبقه‌بند SVM، با معیار MCNemars، روش WFLE بعد از روش NFLE نسبت به سایر روش‌ها برتری دارد.

به طور کلی، می‌توان مشاهده کرد که روش WFLE در بیشتر موارد نسبت به سایر روش‌ها برتری دارد. هنگامی که تعداد نمونه‌های آموزشی محدودی در دسترس است، WFLE روشی مناسب برای استخراج ویژگی است. اما توجه داشته باشید

که استفاده از روش‌های بر مبنای FL از قبیل NFLE و WFLE در تعداد نمونه‌ی آموزشی بالا چندان مناسب نیست. وقتی نمونه‌ی آموزشی واقعی به تعداد زیاد و کافی موجود باشد، استفاده از نمونه‌های آموزشی مجازی دیگر معنا نخواهد داشت و چه بسا منجر به کاهش صحت طبقه بندی گردد. از سوی دیگر، از آنجایی که تعداد نمونه‌های آموزشی مجازی تولید شده در هر دسته برابر $\frac{(n_{tc-1}) \times (n_{tc-2})}{p}$ است، با افزایش تعداد نمونه‌های آموزشی، پیچیدگی روش‌های NFLE و WFLE افزایش خواهد یافت.

مقایسه زمانی روش‌های مختلف استخراج ویژگی برای داده‌ی ایندیانا و با استفاده از طبقه‌بند SVM در جدول ۷ بیان شده‌اند. همان‌طور که از نتایج آزمایش‌ها مشخص است، WFLE نسبت به سایر روش‌ها به زمان محاسباتی بیشتری نیاز دارد و دلیل آن تولید و استفاده از نمونه‌های آموزشی مجازی در فضای باند کامل است. در روش NFLE، ابتدا بعد داده با استفاده از روش PCA کاهش یافته و سپس نمونه‌های مجازی تولید و در محاسبه ماتریس‌های پراکندگی استفاده می‌شوند. اما در روش WFLE، به دلایل توضیح داده شده، کاهش بعد داده اولیه توسط PCA حذف شده و محاسبات با تعداد باند کامل داده‌ی ابرطیفی انجام می‌گیرد که این به مراتب بار محاسباتی را افزون خواهد کرد. پس از WFLE، روش NWFLE که نیاز به محاسبه‌ی میانگین‌های وزن دار دارد، بار محاسباتی زیادی دارد. روش‌هایی چون NPE و LPP که نیازی به تخمین ماتریس‌های پراکندگی ندارند، سریع‌ترین روش‌ها بوده و افزایش تعداد نمونه‌های آموزشی، در زمان محاسباتی آن‌ها تاثیر چندانی ندارد. در حالی که بار محاسباتی روش‌های NWFLE، NFLE و WFLE با

افزایش
تعداد
جدول ۱:
مشخصات
داده
های
ابرطیفی
مورد



استفاده

داده	سنجنده	بازه فرکانسی	دقت مکانی	اندازه تصویر	تعداد کلاس‌ها	تعداد باندهای طیفی (بعد از حذف باندهای نویزی و جذب آب)
ایندیانا	AVIRIS	۰/۴ تا ۲/۵ میکرومتر	۲۰ متر	۱۴۵×۱۴۵	۱۰	۲۰۰
دانشگاه پابوا	ROSIS	۰/۴۳ تا ۰/۸۶ میکرومتر	۱/۳ متر	۳۴۰×۶۱۰	۹	۱۰۳
سالیانس-A	AVIRIS	۰/۴ تا ۲/۵ میکرومتر	۳/۷ متر	۸۶×۸۳	۶	۲۰۴

نمونه‌های آموزشی به شدت افزایش می‌یابد. پس از NPE و LPP، LDA سرعت بیشتری نسبت به سایر روش‌ها دارد چرا که ماتریس‌های پراکندگی را بدون هیچ وزن‌دهی و تنها با نمونه‌های آموزشی اصلی محاسبه می‌کند. افزایش زمان محاسباتی GDA نسبت به LDA به دلیل انجام محاسبات در فضای کرنل می‌باشد.

۵- نتیجه‌گیری

یک روش استخراج ویژگی نظارت شده برای طبقه‌بندی داده‌های ابرطیفی در این مقاله پیشنهاد شد. روش پیشنهادی که WFLE نام گرفت، دارای کارایی خوبی با استفاده از تعداد نمونه‌های آموزشی محدود است. روش WFLE از قدرت نمونه‌های آموزشی مجازی استفاده می‌کند و با وزن‌دهی مناسب به این نمونه‌ها، سعی در اصلاح بیشتر نمونه‌های نامطلوبی دارد که سبب ایجاد خطا در طبقه‌بندی می‌شوند. روش پیشنهادی با تعدادی از روش‌های معروف و پرکاربرد استخراج ویژگی مقایسه گردید و نتایج آزمایش‌های انجام شده با استفاده از نمونه‌های آموزشی محدود، برتری WFLE را نسبت به سایر روش‌ها نشان داد.

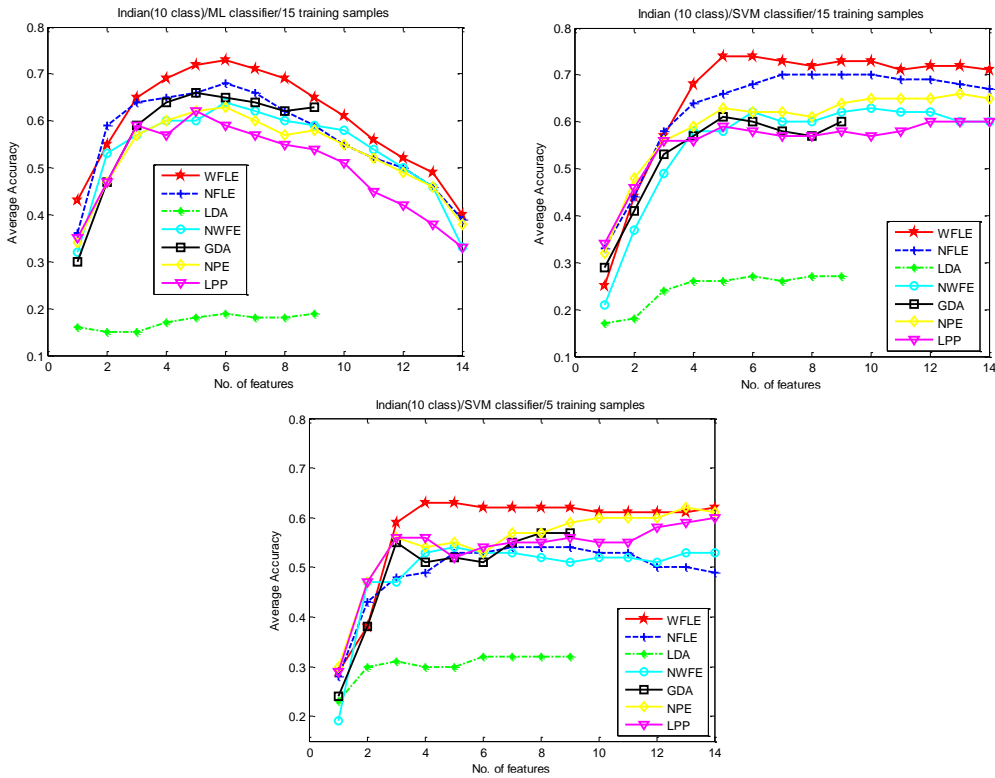
شکل ۵: بلوک‌دیگرام روش پیشنهادی

جدول ۲: بالاترین صحت طبقه‌بندی به دست آمده برای روش‌های مختلف استخراج ویژگی. عدد نوشته شده در پرانتزها، تعداد ویژگی‌هایی را که بالاترین صحت متوسط طبقه‌بندی را فراهم کرده‌اند، نشان می‌دهد. در هر آزمایش، اعداد مربوط به روشی که بالاترین صحت را فراهم کرده‌اند پررنگ شده‌اند.

داده	طبقه‌بند	تعداد نمونه‌های آموزشی	WFLE	NFLE	LDA	NWFE	GDA	NPE	LPP
ایندیانا	SVM	۵	۰.۶۳ (۴)	۰.۵۴ (۷)	۰.۳۲ (۶)	۰.۵۴ (۵)	۰.۵۷ (۸)	۰.۶۲ (۱۳)	۰.۶۰ (۱۴)
		۱۵	۰.۷۴ (۵)	۰.۷۰ (۷)	۰.۲۷ (۶)	۰.۶۳ (۱۰)	۰.۶۱ (۵)	۰.۶۶ (۱۳)	۰.۶۰ (۱۲)
	ML	۱۵	۰.۷۳ (۶)	۰.۶۸ (۶)	۰.۱۹ (۶)	۰.۶۴ (۶)	۰.۶۶ (۵)	۰.۶۳ (۶)	۰.۶۲ (۵)
دانشگاه پلویا	SVM	۵	۰.۷۵ (۵)	۰.۷۱ (۷)	۰.۳۶ (۶)	۰.۷۲ (۵)	۰.۷۵ (۸)	۰.۷۵ (۷)	۰.۷۵ (۱۲)
		۱۵	۰.۸۱ (۱۰)	۰.۸۲ (۷)	۰.۵۴ (۵)	۰.۸۰ (۱۱)	۰.۷۵ (۷)	۰.۸۳ (۱۲)	۰.۸۳ (۱۱)
	ML	۱۵	۰.۸۰ (۴)	۰.۸۰ (۴)	۰.۵۰ (۵)	۰.۷۹ (۵)	۰.۸۰ (۵)	۰.۷۹ (۴)	۰.۷۹ (۶)
سالیانس - A	SVM	۵	۰.۹۵ (۳)	۰.۸۷ (۴)	۰.۶۶ (۳)	۰.۹۰ (۴)	۰.۹۰ (۲)	۰.۹۱ (۵)	۰.۹۱ (۵)
		۱۵	۰.۹۹ (۵)	۰.۹۹ (۶)	۰.۵۰ (۲)	۰.۹۷ (۵)	۰.۹۶ (۵)	۰.۹۸ (۴)	۰.۹۸ (۶)
	ML	۱۵	۰.۹۹ (۳)	۰.۹۹ (۳)	۰.۵۷ (۳)	۰.۹۹ (۴)	۰.۹۹ (۴)	۰.۹۹ (۳)	۰.۹۹ (۳)

جدول ۳: صحت و اعتبار دسته‌ها برای داده ایندیانا با استفاده از ۵ نمونه آموزشی، ۴ ویژگی استخراجی و طبقه‌بند SVM

شماره	نام دسته	تعداد نمونه‌ها	WFLE		NFLE		LDA		NWFE		GDA		NPE		LPP	
			صحت	اعتبار	صحت	اعتبار	صحت	اعتبار	صحت	اعتبار	صحت	اعتبار	صحت	اعتبار	صحت	اعتبار
۱	ذرت-بدون شخم	۱۴۳۴	۰.۴۶	۰.۳۹	۰.۲۵	۰.۲۲	۰.۱۹	۰.۲۶	۰.۳۲	۰.۶۷	۰.۳۹	۰.۲۸	۰.۳۳	۰.۳۶	۰.۲۸	۰.۳۱
۲	ذرت-کم شخم	۸۳۴	۰.۵۱	۰.۴۸	۰.۵۱	۰.۳۰	۰.۴۲	۰.۲۱	۰.۲۹	۰.۵۲	۰.۳۱	۰.۳۷	۰.۳۴	۰.۳۱	۰.۳۹	۰.۳۱
۳	سیزه-درختان	۴۹۷	۰.۹۳	۰.۵۲	۰.۶۳	۰.۴۴	۰.۲۰	۰.۱۷	۰.۷۲	۰.۳۴	۰.۸۵	۰.۴۳	۰.۷۱	۰.۴۷	۰.۷۸	۰.۴۸
۴	سیزه-چمنزار	۷۴۷	۰.۷۶	۰.۷۵	۰.۶۶	۰.۵۰	۰.۵۳	۰.۱۶	۰.۸۱	۰.۵۲	۰.۴۴	۰.۷۱	۰.۳۷	۰.۶۹	۰.۷۶	۰.۷۵
۵	کاه و خاشاک	۴۸۹	۰.۹۹	۱.۰۰	۰.۹۲	۰.۹۷	۰.۵۰	۰.۷۹	۰.۹۷	۱.۰۰	۰.۹۹	۰.۸۵	۰.۹۹	۰.۹۳	۰.۹۹	۰.۹۱
۶	سویا-بدون شخم	۹۶۸	۰.۵۴	۰.۴۷	۰.۱۹	۰.۲۶	۰.۰۶	۰.۱۰	۰.۴۴	۰.۳۶	۰.۵۵	۰.۳۵	۰.۶۵	۰.۴۴	۰.۶۳	۰.۴۲
۷	سویا-کم شخم	۲۴۶۸	۰.۳۶	۰.۵۶	۰.۲۳	۰.۳۵	۰.۱۰	۰.۴۲	۰.۶۹	۰.۵۶	۰.۱۷	۰.۳۴	۰.۴۸	۰.۵۸	۰.۴۵	۰.۵۹
۸	سویا-شخم کامل	۶۱۴	۰.۴۴	۰.۳۱	۰.۲۲	۰.۲۲	۰.۲۵	۰.۱۴	۰.۲۹	۰.۲۸	۰.۲۸	۰.۲۳	۰.۲۲	۰.۲۳	۰.۲۶	۰.۳۰
۹	بیشه	۱۲۹۴	۰.۸۲	۰.۹۶	۰.۶۸	۰.۹۵	۰.۵۲	۰.۶۸	۰.۵۶	۰.۸۶	۰.۷۳	۰.۹۳	۰.۸۲	۰.۸۸	۰.۷۹	۰.۹۲
۱۰	ساختمان-سیزه	۳۸۰	۰.۴۶	۰.۵۴	۰.۶۵	۰.۳۷	۰.۲۰	۰.۲۱	۰.۲۴	۰.۵۱	۰.۴۲	۰.۳۴	۰.۵۴	۰.۳۴	۰.۳۰	۰.۴۵
صحت متوسط و اعتبار متوسط			۰.۶۳	۰.۶۰	۰.۴۹	۰.۴۶	۰.۳۰	۰.۳۱	۰.۵۳	۰.۵۶	۰.۵۱	۰.۴۸	۰.۵۴	۰.۵۲	۰.۵۶	۰.۵۴



شکل ۶: صحت متوسط طبقه‌بندی در برابر تعداد ویژگی‌های استخراج شده برای داده‌ی ایندیانا

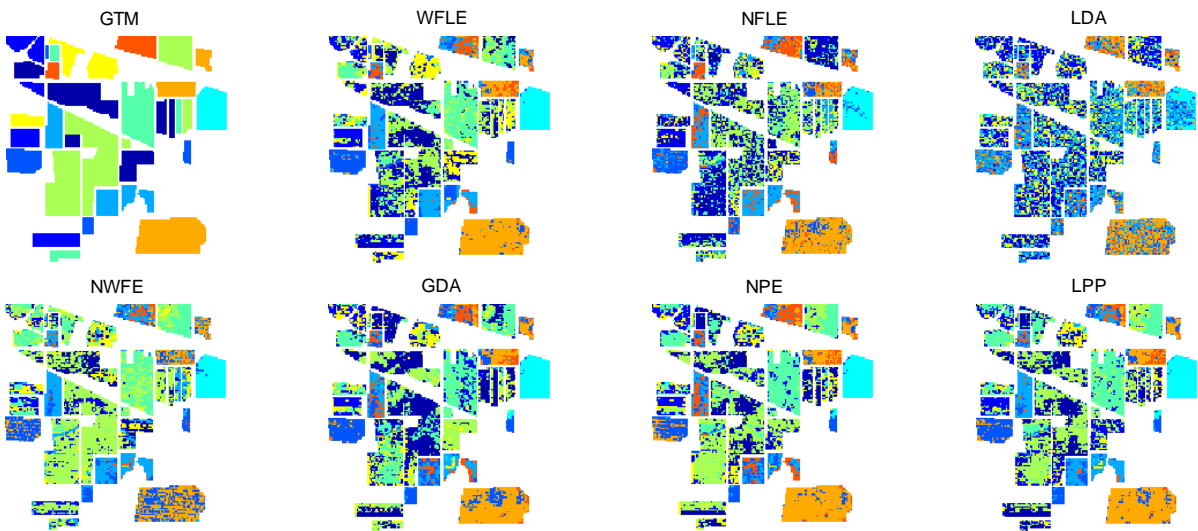
جدول ۴: صحت و اعتبار دسته‌ها برای داده پایویا با استفاده از ۵ نمونه آموزشی، ۵ ویژگی استخراجی و طبقه‌بند SVM

شماره	نام دسته	تعداد نمونه‌ها	WFLE		NFLE		LDA		NWFE		GDA		NPE		LPP	
			صحت	اعتبار	صحت	اعتبار	صحت	اعتبار	صحت	اعتبار	صحت	اعتبار	صحت	اعتبار	صحت	اعتبار
۱	آسفالت	۶۶۳۱	۰.۵۸	۰.۸۷	۰.۵۲	۰.۸۹	۰.۳۶	۰.۱۴	۰.۵۹	۰.۸۳	۰.۷۱	۰.۸۸	۰.۶۴	۰.۸۸	۰.۵۶	۰.۸۷
۲	چمن‌زار	۱۸۶۴ ۹	۰.۵۸	۰.۸۵	۰.۵۵	۰.۸۷	۰.۳۶	۰.۷۹	۰.۲۷	۰.۷۷	۰.۶۸	۰.۸۱	۰.۳۵	۰.۸۱	۰.۵۷	۰.۸۳
۳	شن و سنگریزه	۲۰۹۹	۰.۶۲	۰.۳۴	۰.۵۰	۰.۴۵	۰.۳۰	۰.۲۰	۰.۳۶	۰.۴۱	۰.۵۳	۰.۵۹	۰.۳۸	۰.۴۱	۰.۴۶	۰.۳۳
۴	درختان	۳۰۶۴	۰.۹۱	۰.۵۵	۰.۸۴	۰.۷۱	۰.۷۴	۰.۷۲	۰.۹۵	۰.۴۸	۰.۹۶	۰.۴۵	۰.۹۷	۰.۴۸	۰.۹۷	۰.۴۷
۵	صفحات فلزی	۱۳۴۵	۰.۹۹	۱.۰۰	۱.۰۰	۱.۰۰	۰.۶۸	۰.۸۵	۰.۹۸	۰.۹۸	۰.۹۹	۰.۵۹	۰.۹۹	۰.۹۶	۰.۹۹	۰.۹۶
۶	زمین دست نخورده	۵۰۲۹	۰.۵۷	۰.۳۳	۰.۵۸	۰.۲۴	۰.۱۲	۰.۲۰	۰.۷۰	۰.۲۴	۰.۲۹	۰.۳۷	۰.۶۲	۰.۲۶	۰.۴۷	۰.۳۳
۷	قیر	۱۳۳۰	۰.۸۵	۰.۴۵	۰.۵۷	۰.۳۸	۰.۱۷	۰.۱۰	۰.۸۱	۰.۴۸	۰.۶۵	۰.۴۶	۰.۷۱	۰.۴۷	۰.۷۰	۰.۴۶
۸	آجر	۳۶۸۲	۰.۶۳	۰.۷۱	۰.۶۸	۰.۵۱	۰.۰۸	۰.۲۸	۰.۷۹	۰.۶۵	۰.۸۲	۰.۶۹	۰.۸۸	۰.۶۳	۰.۸۶	۰.۶۴
۹	سایه	۹۴۷	۱.۰۰	۱.۰۰	۰.۹۹	۱.۰۰	۰.۳۸	۰.۰۸	۰.۹۹	۰.۸۴	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰
صحت متوسط و اعتبار متوسط			۰.۷۵	۰.۶۸	۰.۶۹	۰.۶۷	۰.۳۵	۰.۳۷	۰.۷۲	۰.۶۳	۰.۷۴	۰.۶۵	۰.۷۳	۰.۶۵	۰.۷۳	۰.۶۵

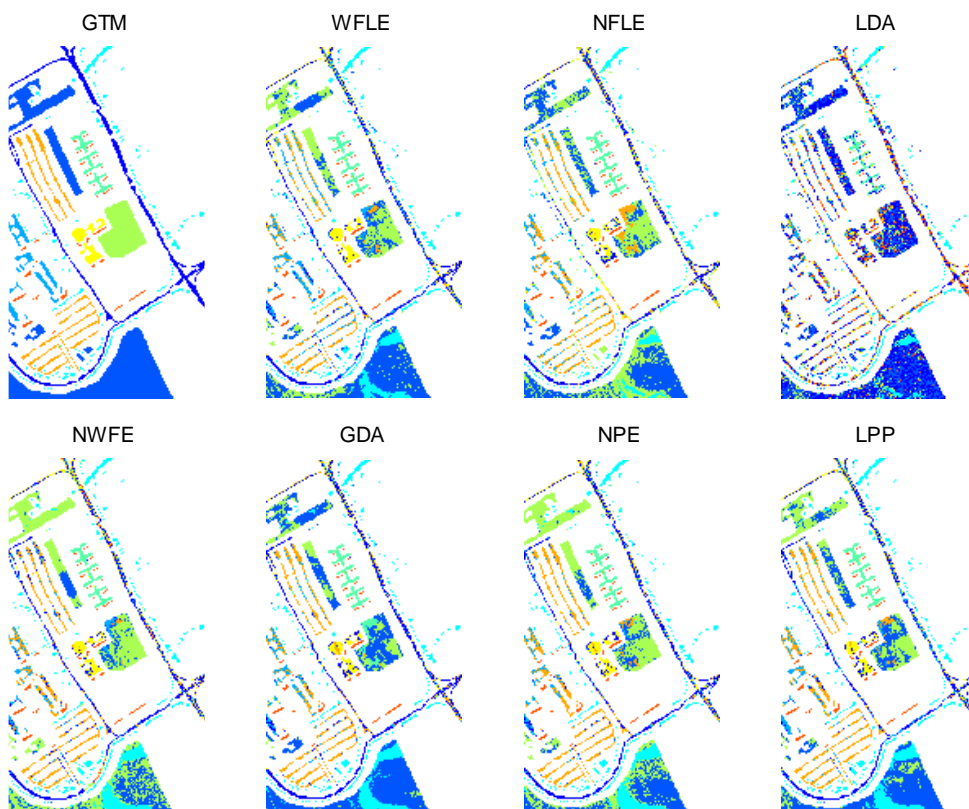


جدول ۵: صحت و اعتبار دسته‌ها برای داده سالیانس-A با استفاده از ۵ نمونه آموزشی، ۳ ویژگی استخراجی و طبقه‌بند SVM

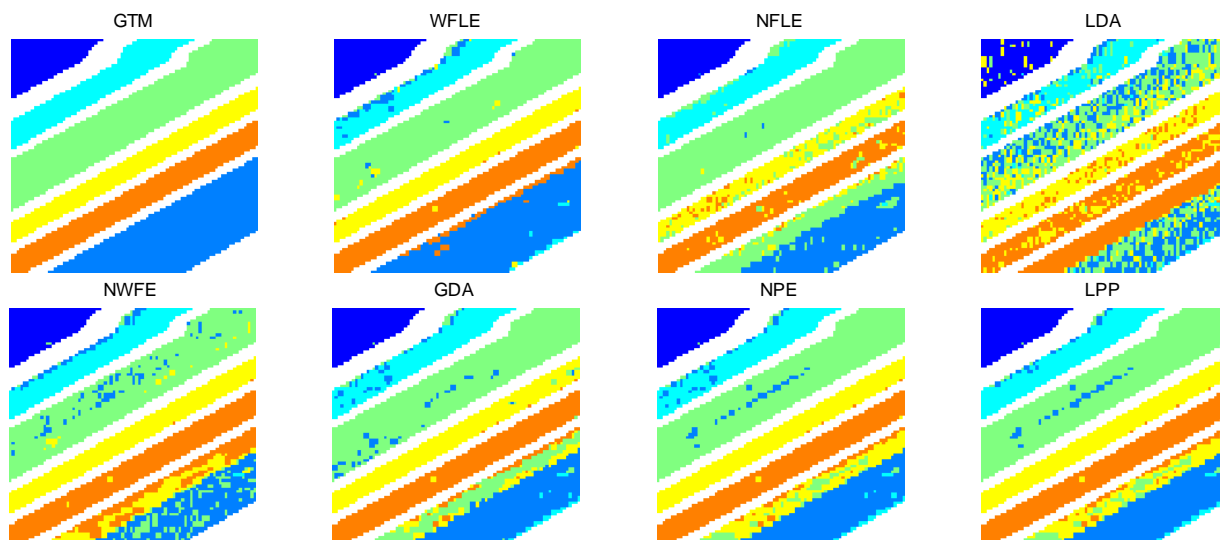
دسته			WFLE		NFLE		LDA		NWFE		GDA		NPE		LPP	
شماره	نام دسته	تعداد نمونه‌ها	صحت	اعتبار	صحت	اعتبار	صحت	اعتبار	صحت	اعتبار	صحت	اعتبار	صحت	اعتبار	صحت	اعتبار
۱	بروکلی-سبزه- علف هرز ۱	۳۹۱	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰	۰.۸۸	۱.۰۰	۰.۹۹	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰
۲	ذرت-سبزه پیر-علف هرز	۱۳۴۳	۰.۸۹	۰.۹۳	۰.۵۶	۰.۹۹	۰.۲۸	۰.۴۴	۰.۴۳	۰.۷۸	۰.۵۷	۰.۸۷	۰.۵۶	۰.۸۷	۰.۵۶	۰.۸۹
۳	کاهو-۴ هفته	۶۱۶	۰.۸۲	۰.۹۵	۰.۸۹	۱.۰۰	۰.۶۹	۰.۹۰	۰.۸۷	۰.۹۹	۰.۸۸	۰.۹۳	۰.۸۸	۰.۹۳	۰.۹۱	۰.۹۵
۴	کاهو-۵ هفته	۱۵۲۵	۰.۹۸	۰.۹۸	۱.۰۰	۰.۶۹	۰.۴۹	۰.۶۴	۰.۹۲	۰.۸۸	۰.۹۶	۰.۸۰	۰.۹۶	۰.۹۱	۰.۹۶	۰.۹۱
۵	کاهو-۶ هفته	۶۷۴	۰.۹۹	۰.۹۵	۰.۷۶	۰.۸۱	۰.۸۳	۰.۴۵	۱.۰۰	۰.۶۸	۰.۹۷	۰.۸۲	۰.۹۹	۰.۶۵	۰.۹۹	۰.۶۶
۶	کاهو-۷ هفته	۷۹۹	۰.۹۹	۰.۸۷	۰.۹۴	۰.۹۳	۰.۷۸	۰.۴۹	۰.۹۹	۰.۷۲	۰.۹۹	۰.۹۰	۰.۹۹	۰.۹۰	۰.۹۹	۰.۸۹
صحت متوسط و اعتبار متوسط			۰.۹۵	۰.۹۵	۰.۸۶	۰.۹۰	۰.۶۶	۰.۶۵	۰.۸۷	۰.۸۴	۰.۸۹	۰.۸۹	۰.۹۰	۰.۸۸	۰.۹۰	۰.۸۸



شکل ۷: نقشه مرجع طبقه‌بندی (GTM) و نقشه‌های طبقه‌بندی به دست آمده با استفاده از روش‌های مختلف استخراج ویژگی WFLE، NFLE، LDA، NWFE، GDA، NPE و LPP برای داده ایندیانا با استفاده از ۵ نمونه آموزشی، ۴ ویژگی استخراجی و طبقه‌بند SVM



شکل ۸: نقشه مرجع طبقه‌بندی (GTM) و نقشه‌های طبقه‌بندی به دست آمده با استفاده از روش‌های مختلف استخراج ویژگی WFLE، NFLE، LDA، NWFE، GDA، NPE و LPP برای داده پویا با استفاده از ۵ نمونه آموزشی، ۵ ویژگی استخراجی و طبقه‌بند SVM



شکل ۹: نقشه مرجع طبقه‌بندی (GTM) و نقشه‌های طبقه‌بندی به دست آمده با استفاده از روش‌های مختلف استخراج ویژگی WFLE، NFLE، LDA، NWFE، GDA، NPE و LPP برای داده سالیانس-A با استفاده از ۵ نمونه آموزشی، ۳ ویژگی استخراجی و طبقه‌بند SVM

جدول ۶: نتایج آزمون McNemars. هر خانه‌ی جدول بیان‌کننده Z_{rc} است که r شماره سطر و c شماره ستون جدول است.

داده ایندیانا، طبقه‌بند SVM، ۵ نمونه آموزشی و ۴ ویژگی استخراجی

	WFLE	NFLE	LDA	NWFE	GDA	NPE	LPP
WFLE	۰	۲۵.۵۰	۴۴.۲۳	۴.۴۴	۲۲.۹۱	۸.۰۶	۶.۵۹
NFLE	-۲۵.۵۰	۰	۲۵.۶۹	-۱۸.۵۱	-۴.۰۳	-۱۶.۴۰	-۱۸.۳۸
LDA	-۴۴.۲۳	-۲۵.۶۹	۰	-۳۹.۲۰	-۲۷.۰۴	-۳۶.۷۹	-۳۸.۷۳
NWFE	-۴.۴۴	۱۸.۵۱	۳۹.۲۰	۰	۱۵.۶۲	۳.۱۱	۱.۳۵
GDA	-۲۲.۹۱	۴.۰۳	۲۷.۰۴	-۱۵.۶۲	۰	-۱۶.۸۶	-۱۷.۷۳
NPE	-۸.۰۶	۱۶.۴۰	۳۶.۷۹	-۳.۱۱	۱۶.۸۶	۰	-۲.۶۵
LPP	-۶.۵۹	۱۸.۳۸	۳۸.۷۳	-۱.۳۵	۱۷.۷۳	۲.۶۵	۰

داده ایندیانا، طبقه‌بند SVM، ۱۵ نمونه آموزشی و ۵ ویژگی استخراجی

	WFLE	NFLE	LDA	NWFE	GDA	NPE	LPP
WFLE	۰	۲۱.۳۰	۶۱.۷۳	۲۴.۷۴	۲۲.۶۹	۱۹.۸۰	۲۴.۰۷
NFLE	-۲۱.۳۰	۰	۵۱.۱۳	۷.۶۵	۳.۷۵	۰.۱۱	۵.۳۱
LDA	-۶۱.۷۳	-۵۱.۱۳	۰	-۴۶.۲۳	-۴۹.۲۵	-۵۰.۵۸	-۴۶.۸۴
NWFE	-۲۴.۷۴	-۷.۶۵	۴۶.۲۳	۰	-۴.۴۱	-۸.۱۶	-۲.۴۲
GDA	-۲۲.۶۹	-۳.۷۵	۴۹.۲۵	۴.۴۱	۰	-۴.۱۲	۲.۰۹
NPE	-۱۹.۸۰	۰.۱۱	۵۰.۵۸	۸.۱۶	۴.۱۲	۰	۸.۵۸
LPP	-۲۴.۰۷	-۵.۳۱	۴۶.۸۴	۲.۴۲	-۲.۰۹	-۸.۵۸	۰

داده ایندیانا، طبقه‌بند ML، ۱۵ نمونه آموزشی و ۶ ویژگی استخراجی

	WFLE	NFLE	LDA	NWFE	GDA	NPE	LPP
WFLE	۰	۱۴.۰۳	۶۱.۷۲	۱۷.۴۳	۱۵.۱۴	۱۹.۰۲	۲۴.۳۵
NFLE	-۱۴.۰۳	۰	۵۵.۲۷	۴.۸۳	۱.۹۶	۷.۰۷	۱۳.۰۳
LDA	-۶۱.۷۲	-۵۵.۲۷	۰	-۵۲.۸۶	-۵۴.۸۷	-۵۱.۶۷	-۴۷.۵۷
NWFE	-۱۷.۴۳	-۴.۸۳	۵۲.۸۶	۰	-۲.۹۴	۲.۸۴	۹.۹۴
GDA	-۱۵.۱۴	-۱.۹۶	۵۴.۸۷	۲.۹۴	۰	۵.۶۶	۱۲.۰۷
NPE	-۱۹.۰۲	۷.۰۷	۵۱.۶۷	-۲.۸۴	-۵.۶۶	۰	۹.۰۴
LPP	-۲۴.۳۵	-۱۳.۰۳	۴۷.۵۷	-۹.۹۴	-۱۲.۰۷	-۹.۰۴	۰

داده پاپوا، طبقه‌بند SVM، ۵ نمونه آموزشی و ۵ ویژگی استخراجی

	WFLE	NFLE	LDA	NWFE	GDA	NPE	LPP
WFLE	۰	۱۲.۶۲	۸۸.۹۰	۴۴.۰۱	-۱.۴۷	۲۸.۵۶	۳.۳۸
NFLE	-۱۲.۶۲	۰	۸۳.۶۱	۳۰.۶۲	-۲۸.۷۷	۱۴.۴۱	-۱۰.۲۷
LDA	-۸۸.۹۰	-۸۳.۶۱	۰	-۵۶.۸۰	-۱۰۱.۶۸	-۶۸.۷۹	-۸۶.۳۱
NWFE	-۴۴.۰۱	۳۰.۶۲	۵۶.۸۰	۰	-۵۶.۹۰	-۲۲.۹۸	-۴۲.۶۳
GDA	۱.۴۷	۲۸.۷۷	۱۰۱.۶۸	۵۶.۹۰	۰	۴۵.۹۸	۲۳.۸۴
NPE	-۲۸.۵۶	-۱۴.۴۱	۶۸.۷۹	۲۲.۹۸	-۴۵.۹۸	۰	-۲۹.۹۶
LPP	-۳.۳۸	۱۰.۲۷	۸۶.۳۱	۴۲.۶۳	-۲۳.۸۴	۲۹.۹۶	۰

داده پاپوا، طبقه‌بند SVM، ۱۵ نمونه آموزشی و ۵ ویژگی استخراجی

	WFLE	NFLE	LDA	NWFE	GDA	NPE	LPP
WFLE	۰	-۲.۹۸	۵۲.۴۸	۱۵.۲۱	۳۰.۳۱	۱۵.۹۰	۴۰.۳۷
NFLE	۲.۹۸	۰	۵۳.۷۵	۱۶.۵۹	۳۲.۳۶	۱۶.۲۹	۳۸.۰۹
LDA	-۵۲.۴۸	-۵۳.۷۵	۰	-۳۹.۴۳	-۱۹.۲۱	-۴۰.۶۶	-۲۰.۴۰
NWFE	-۱۵.۲۱	-۱۶.۵۹	۳۹.۴۳	۰	۱۸.۴۲	-۱.۰۹	۲۴.۲۵
GDA	-۳۰.۳۱	-۳۲.۳۶	۱۹.۲۱	-۱۸.۴۲	۰	-۱۹.۶۵	-۰.۳۳
NPE	-۱۵.۹۰	-۱۶.۲۹	۴۰.۶۶	۱.۰۹	۱۹.۶۵	۰	۳۰.۲۰
LPP	-۴۰.۳۷	-۳۸.۰۹	۲۰.۴۰	-۲۴.۲۵	۰.۳۳	-۳۰.۲۰	۰

ادامه جدول ۶: نتایج آزمون McNemars. هر خانه‌ی جدول بیان‌کننده Z_{rc} است که r شماره سطر و c شماره ستون جدول است.

داده پاپویا، طبقه‌بند ML، ۱۵ نمونه آموزشی و ۴ ویژگی استخراجی

	WFLE	NFLE	LDA	NWFE	GDA	NPE	LPP
WFLE	۰	۸.۹۵	۶۹.۴۴	۱۳.۱۸	۸.۷۲	۴.۷۷	۷.۳۹
NFLE	-۸.۹۵	۰	۶۴.۶۳	۵.۷۳	۱.۶۷	-۲.۰۳	۰.۴۲
LDA	-۶۹.۴۴	-۶۴.۶۳	۰	-۵۸.۵۹	-۶۰.۶۰	-۶۳.۵۰	-۶۱.۵۸
NWFE	-۱۳.۱۸	-۵.۷۳	۵۸.۵۹	۰	-۵.۸۵	-۱۰.۷۷	-۷.۳۱
GDA	-۸.۷۲	-۱.۶۷	۶۰.۶۰	۵.۸۵	۰	-۱۱.۶۸	-۳.۶۶
NPE	-۴.۷۷	۲.۰۳	۶۳.۵۰	۱۰.۷۷	۱۱.۶۸	۰	۹.۱۵
LPP	-۷.۳۹	-۰.۴۲	۶۱.۵۸	۷.۳۱	۳.۶۶	-۹.۱۵	۰

داده سالیانس-A، طبقه‌بند SVM، ۵ نمونه آموزشی و ۳ ویژگی استخراجی

	WFLE	NFLE	LDA	NWFE	GDA	NPE	LPP
WFLE	۰	۲۰.۶۶	۴۳.۴۴	۲۳.۸۴	۱۸.۷۸	۱۸.۱۸	۱۷.۱۳
NFLE	-۲۰.۶۶	۰	۳۳.۵۹	۴.۶۸	-۶.۳۳	-۶.۶۴	-۸.۴۹
LDA	-۴۳.۴۴	-۳۳.۵۹	۰	-۳۲.۴۴	-۳۷.۰۸	-۳۷.۵۷	-۳۸.۵۰
NWFE	-۲۳.۸۴	-۴.۶۸	۳۲.۴۴	۰	-۱۰.۹۲	-۱۱.۸۹	-۱۳.۸۶
GDA	-۱۸.۷۸	۶.۳۳	۳۷.۰۸	۱۰.۹۲	۰	-۰.۴۷	-۳.۹۴
NPE	-۱۸.۱۸	۶.۶۴	۳۷.۵۷	۱۱.۸۹	۰.۴۷	۰	-۴.۹۶
LPP	-۱۷.۱۳	۸.۴۹	۳۸.۵۰	۱۳.۸۶	۳.۰۴	۴.۹۶	۰

داده سالیانس-A، طبقه‌بند SVM، ۱۵ نمونه آموزشی و ۴ ویژگی استخراجی

	WFLE	NFLE	LDA	NWFE	GDA	NPE	LPP
WFLE	۰	۴.۰۹	۵۳.۴۵	۹.۳۸	۱۵.۴۱	۱.۹۰	۰.۵۲
NFLE	-۴.۰۹	۰	۵۳.۳۰	۵.۸۴	۱۱.۸۶	-۲.۲۲	-۳.۵۹
LDA	-۵۳.۴۵	-۵۳.۳۰	۰	-۵۱.۱۰	-۴۸.۰۵	-۵۳.۴۴	-۵۳.۶۹
NWFE	-۹.۳۸	-۵.۸۴	۵۱.۱۰	۰	۶.۸۶	-۸.۷۳	-۹.۹۴
GDA	-۱۵.۴۱	-۱۱.۸۶	۴۸.۰۵	-۶.۸۶	۰	-۱۶.۴۱	-۱۶.۸۷
NPE	-۱.۹۰	۲.۲۲	۵۳.۴۴	۸.۷۳	۱۶.۴۱	۰	-۲.۷۹
LPP	-۰.۵۲	۳.۵۹	۵۳.۶۹	۹.۹۴	۱۶.۸۷	۲.۷۹	۰

داده سالیانس-A، طبقه‌بند ML، ۱۵ نمونه آموزشی و ۳ ویژگی استخراجی

	WFLE	NFLE	LDA	NWFE	GDA	NPE	LPP
WFLE	۰	۴.۱۰	۵۰.۶۹	۲.۵۹	۱.۸۹	۴.۰۰	۲.۶۰
NFLE	-۴.۱۰	۰	۵۰.۲۰	-۱.۵۰	-۲.۲۲	۰.۰۹	-۱.۶۱
LDA	-۵۰.۶۹	-۵۰.۲۰	۰	-۵۰.۴۳	-۵۰.۵۶	-۵۰.۱۳	-۵۰.۳۱
NWFE	-۲.۵۹	۱.۵۰	۵۰.۴۳	۰	-۰.۷۶	۱.۶۳	-۰.۱۰
GDA	-۱.۸۹	۲.۲۲	۵۰.۵۶	۰.۷۶	۰	۳.۶۲	۰.۸۰
NPE	-۴.۰۰	-۰.۰۹	۵۰.۱۳	-۱.۶۳	-۳.۶۲	۰	-۲.۰۴
LPP	-۲.۶۰	۱.۶۱	۵۰.۳۱	۰.۱۰	-۰.۸۰	۲.۰۴	۰

فرآیند استخراج
مختلف با استفاده از
داده‌ی ایندیانا

	تعداد نمونه‌های آموزشی	WFLE	NFLE	LDA	NWFE	GDA	NPE	LPP
		۵	۲.۸۱	۰.۵۵	۰.۶۷	۲.۴۳	۰.۷۰	۰.۴۷
زمین (ثانیه)	۱۵	۲۵۲.۲۷	۱۶.۰۰	۰.۷۳	۷۷.۵۴	۰.۸۱	۰.۵۹	۰.۴۹

جدول ۷: مقایسه زمانی
ویژگی در روش‌های
طبقه‌بند SVM برای



مراجع

- classification and retrieval," IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 11, pp. 1335-1339, Nov. 2000.
- [18] Y. W. Pang, Y. Yuan, and X. Li, "Generalized nearest feature line for subspace learning," Electron. Lett., vol. 43, no. 20, pp. 1079-1080, Sep. 2007.
- [19] J. Lu and Y. P. Tan, "Uncorrelated discriminant nearest feature line analysis for face recognition," IEEE Signal Process. Lett., vol. 17, no. 2, pp. 185-188, Feb. 2010.
- [20] S. Theodoridis and K. Koutroumbas, Pattern Recognition, 4th Ed., Elsevier Academic Press, San Diego, California, pp. 119-127, 2009.
- [21] S. A. Hosseini and H. Ghassemian, "A new fast algorithm for multiclass hyperspectral image classification with SVM," International Journal of Remote Sensing, vol. 32, no. 23, pp. 8657-8683, Dec. 2011.
- [22] C. Chang and C. Linin, LIBSVM—A Library for Support Vector Machines [Online]. Available: <http://www.csie.ntu.edu.tw/~cjlin/libsvm>, 2008.
- [23] 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] ر. کامران، ح. نظام آبادی و س. یزدی، ترمیم تصاویر رنگی با نواحی مخدوش بزرگ بر اساس تجزیه تصویر به مولفه‌های بافت و ساختار، نشریه مهندسی برق و الکترونیک ایران، پاییز و زمستان 1390؛ 8 (2): 13-24.
- [2] G. F. Hughes, "On the mean accuracy of statistical pattern recognition," IEEE Trans. Inf. Theory, vol. IT-14, no. 1, pp. 55-63, Jan. 1968.
- [3] S. Li, H. Wu, D. Wan, J. Zhu, "An effective feature selection method for hyperspectral image classification based on genetic algorithm and support vector machine," Knowledge-Based Systems, vol. 24, pp. 40-48, 2011.
- [4] J. Yin, Y. Wang, and J. Hu, "A new dimensionality reduction algorithm for hyperspectral image data using evolutionary strategy," IEEE Trans. Ind. Inf., vol. 8, no. 4, pp. 935-943, Nov. 2012.
- [5] 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.
- [6] م. مهدی‌خانی و م.ح. کهایبی، معرفی الگوریتم جدید LICAD برای حل مشکل جایگشت محلی الگوریتم ICA، نشریه مهندسی برق و الکترونیک ایران، پاییز و زمستان 1393؛ 11 (2): 25-32.
- [7] J. Yin, C. Gao, and X. Jia, "Wavelet Packet Analysis and Gray Model for Feature Extraction of Hyperspectral Data," IEEE Geoscience and Remote Sensing Lett., vol. 10, no. 4, pp. 682-686, July 2013.
- [8] H. Ghassemian and D.A. Landgrebe, "Object-Oriented Feature Extraction Method for Image Data Compaction," IEEE Control Systems Magazine, vol. 8, no. 3, pp. 42-48, June 1988.
- [9] 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, March 2013.
- [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.
- [13] X. F. He, D. Cai, S. C. Yan, and H. J. Zhang, "Neighborhood preserving embedding," in Proc. 10th IEEE Int. Conf. Comput. Vis., vol. 2, pp. 1208-1213, 2005.
- [14] X. F. He and P. Niyogi, "Locality preserving projections," in Proc. Adv. Neural Inf. Process. Syst., vol. 16, pp. 152-160, 2004.
- [15] Y.-L. Chang, J.-N. Liu, C.-C. Han, and Y.-N. Chen, "Hyperspectral Image Classification Using Nearest Feature Line Embedding Approach," IEEE Trans. Geosci. Remote Sens., vol. 52, no. 1, pp. 278-287, Jan. 2014.
- [16] S. Z. Li and J. Lu, "Face recognition using the nearest feature line method," IEEE Trans. Neural Netw., vol. 10, no. 2, pp. 439-443, March 1999.
- [17] S. Z. Li, K. L. Chan, and C. L. Wang, "Performance evaluation of the nearest feature line method in image

- ¹ Class
- ² Weighted Feature Line Embedding
- ³ Feature Line
- ⁴ Linear Discriminant Analysis
- ⁵ Generalized Discriminant Analysis
- ⁶ Nonparametric Weighted Feature Extraction
- ⁷ Neighborhood Preserving Embedding
- ⁸ Locality Preserving Projection
- ⁹ Nearest Feature Line Embedding
- ¹⁰ manifold learning
- ¹¹ adjacency graph
- ¹² similarity matrix
- ¹³ Nearest Feature Line
- ¹⁴ feature line distance
- ¹⁵ position parameter
- ¹⁶ Principal Component Analysis
- ¹⁷ regularization
- ¹⁸ Indian
- ¹⁹ University of Pavia
- ²⁰ Salinas-A
- ²¹ Airborne Visible/Infrared Imaging Spectrometer
- ²² Reflective Optics System Imaging Spectrometer
- ²³ Support Vector Machine
- ²⁴ Maximum Likelihood

