

مقایسه الگوریتم‌های طبقه‌بندی بر روی تصاویر ماهواره‌ای سنجش از دور

نجمه چراغی شیرازی^{۱*}، روزبه حمزه ثیان^۲، اشکان معصومی^۳

۱- دانشگاه آزاد اسلامی واحد بوشهر، عضو هیات علمی گروه مهندسی برق، بوشهر، ایران Nch_shirazi@yahoo.com

۲- دانشگاه آزاد اسلامی واحد بوشهر، عضو هیات علمی گروه مهندسی برق، بوشهر، ایران R_hamzehyan@yahoo.com

۳- دانشگاه آزاد اسلامی واحد اهرم، عضو هیات علمی گروه مهندسی برق، بوشهر، ایران Ashkan.masoomi@yahoo.com

تاریخ دریافت: تاریخ پذیرش:

چکیده:

اگرچه تعداد بسیار زیادی الگوریتم طبقه‌بندی برای تصاویر ارائه شده، اما به ندرت بر روی یک مورد یکسان بایکدیگر مقایسه شده‌اند. در این مقاله، تصاویر ماهواره‌های سنجش از دور با استفاده از دو روش الگوریتم طبقه‌بندی بدون نظارت و هشت الگوریتم طبقه‌بندی با نظارت که شامل تعدادی از الگوریتم‌های رایج طی بیست سال اخیر است، آزموده شدند. تحلیل ما بر روی تصاویر ماهواره‌ای ۱۲ طیفی متمرکز است. در مقایسه الگوریتم‌ها تعداد نمونه آموزشی یکسان فرض شده‌است. الگوریتم‌ها از نظر پیچیدگی، میزان صحت و اعتبار بایکدیگر مقایسه شده‌اند. نتایج نشان می‌دهد که صحت طبقه‌بندی، نسبت مستقیم با تعداد نمونه‌های آموزشی دارد و همچنین کاربرد می‌تواند بسته به اهمیت هریک از پارامترهای فوق الگوریتم کارآمدتر را برگزید.

کلیدواژه‌ها: الگوریتم‌های طبقه‌بندی، سنجش از دور، نمونه آموزشی، نمونه آزمایشی

مقدمه

زمینه بینایی ماشین را معرفی کرده‌اند. همچنین در سال ۲۰۰۳، J. B. Mena در حدود ۲۵۰ روش در زمینه استخراج راه‌ها را معرفی نمود. در کاربردهای سنجش از دور شیوه‌های طبقه‌بندی مختلفی به منظور استخراج داده‌ها معرفی شده‌است [2].

در دو دهه اخیر، روش‌های زیادی به منظور دسته‌بندی داده‌های سنجش از دور شامل روش‌های پارامتریک (برای نمونه طبقه‌بند حداکثر شباهت) و روش‌های غیر پارامتریک (مانند طبقه‌بند نزدیک‌ترین همسایگی) بکار رفتند. هریک از این الگوریتم‌های طبقه‌بندی دارای مشکلات مشخصی هستند که استفاده از آنها را محدود می‌کند. رایج‌ترین روش دسته‌بندی آماری مبتنی بر حداکثر شباهت در سال ۱۹۷۸ توسط Swain و Davis بر روی تصاویر چند طیفی معرفی شد. این روش قدیمی در صورتیکه داده استفاده شده توزیع نرمال نداشته باشد عملکرد خوبی ندارد. برای حل این مسأله سیستم‌های هوش مصنوعی با بکارگیری آنالیز ویژگی معرفی شدند. این روش‌ها

از زمان بکارگیری اولین ماهواره‌های سنجش از دور، در سال ۱۹۷۲، پیشرفت‌های قابل ملاحظه‌ای در فناوری سنسورهای بکار رفته در ماهواره‌های فعال انجام شده‌است. تفکیک پذیری فاصله‌ای تصاویر ماهواره‌ای از ۸۰ به ۰/۴۱ متر بهبود یافته‌است. همچنین تعداد طیف‌ها از چند باند طیفی محدود به چند صد باند طیفی افزایش یافته. بدنبال آن طبقه‌بندی سطح زمین و زمین به منظور استخراج داده‌های حیاتی توسعه پیدا کرده است [1].

استخراج خودکار اطلاعات مکانی از تصاویر ماهواره‌ای موضوع طیف گسترده‌ای از تحقیقات علوم مختلف در چند دهه‌ی اخیر بوده‌است. از سال ۱۹۸۲ که Marr و گروهش، روش‌های بینایی‌ماشین را معرفی کردند، تا به امروز کارهای بسیاری در آن خصوص انجام گرفته‌است. به‌عنوان نمونه در سال ۲۰۰۰، Rosenfeld و گروهش در حدود ۱۷۰۰ مرجع در

چالش اصلی جمع‌آوری داده‌ها هزینه بر بودن آن و عدم آگاهی از تعداد نمونه‌ها برای انجام یک طبقه‌بندی مناسب است [3]. در بخش انتخاب ویژگی، ویژگی‌های بکارگرفته شده برای نمونه‌های با کلاس‌های یکسان باید مشابه و برای کلاس‌های متفاوت باید جدا کننده باشد و همچنین استخراج ویژگی‌ها باید ساده و ویژگی‌های انتخابی باید مقاوم در برابر نویز باشند. در طبقه‌بندی‌های با نظارت دانش پیشین که از نمونه‌های آموزشی بدست می‌آید در کارایی الگوریتم بسیار تعیین کننده است. انتخاب مدل طبقه‌بندی از مهم‌ترین بخش‌های الگوریتم می‌باشد. انتخاب نوع طبقه‌بند و آگاهی از اینکه برای هر مسأله چه طبقه‌بندی باید استفاده کرد، از چالش‌های پیشروی انتخاب مدل است. بخش آموزش مختص الگوریتم‌های بانظارت است.

جدول ۱- مشخصات باندهای طیفی تصویر

شماره طیف	باند طیف برحسب میکرومتر	
	حداکثر	حداقل
۱	۰/۴۹	۰/۴۶
۲	۰/۵۱	۰/۴۸
۳	۰/۵۴	۰/۵۰
۴	۰/۵۷	۰/۵۲
۵	۰/۶۰	۰/۵۴
۶	۰/۶۵	۰/۵۸
۷	۰/۷۰	۰/۶۱
۸	۰/۹۲	۰/۷۲
۹	۱/۴۰	۱/۰۰
۱۰	۱/۸۰	۱/۵۰
۱۱	۲/۶۰	۲/۰۰
۱۲	۱۱/۷۰	۹/۳۰

در این بخش فرایند استفاده داده‌های آموزشی برای به دست آوردن تخمین پارامترهای اولیه طبقه‌بند استفاده می‌شود. کارایی سیستم در بخش ارزیابی اندازه‌گیری شده و نیازهای بهبود اجزاء سیستم تعیین می‌گردد. چگونگی تنظیم پیچیدگی سیستم برای اجتناب از حساس شدن بیش از حد به داده‌های ورودی تکراری در ارزیابی به دست می‌آید. در ادامه برخی از

در سال‌های ۱۹۸۶ تا ۱۹۸۹ گسترش داده شدند که تمامی آنها به اطلاعات پیشین تصویر نیاز دارند. به منظور بهبود تشخیص ویژگی در تصاویر سنجش از دور، بکارگیری شبکه‌های عصبی معرفی شدند. در این روش‌ها کاربر بایستی تعدادی از پارامترها مانند معماری شبکه، نرخ آموزش، تعداد تکرارها و الگوریتم آموزش را تعیین کند. انتخاب هریک از پارامترهای فوق بر عملکرد روش تأثیر بسزایی دارد [3]. نتایج حاصل از طبقه بندی به چند عامل بستگی دارد: مدل طبقه‌بند، انتخاب نمونه آموزشی، پیش پردازش داده‌ها، الگوریتم طبقه‌بندی، مجموعه نمونه‌های آموزشی و آزمایشی، و روش ارزیابی [3]. هدف از این مقاله مقایسه عملکرد هجده الگوریتم طبقه‌بندی بر روی نمونه تصاویر سنجش از دور می‌باشد. پوشش زمینی تصاویر انتخابی دارای پیچیدگی لازم بوده و از الگوهای مناسب برای مقایسه الگوریتم‌های طبقه‌بندی برخوردار می‌باشد. در ادامه ابتدا تصویر چند طیفی معرفی و سپس الگوریتم‌های طبقه‌بندی با نظارت و بدون نظارت به کار رفته به اختصار بیان می‌شوند. نتایج شبیه‌سازی شده بحث و نتیجه‌گیری خواهد شد.

مشخصات تصویر و داده‌ها

تصویر ماهواره‌ای چند طیفی دارای ۱۲ باند طیفی است. تصویر هشت دسته پوشش گیاهی متفاوت را شامل می‌شود. جدول ۱ مشخصات باند طیفی تصویر را نشان می‌دهد [4].

الگوریتم‌های طبقه‌بندی

تمامی الگوریتم‌های طبقه بندی به طور کلی شامل گام‌های دریافت داده، استخراج ویژگی، انتخاب مدل و ارزیابی می‌باشند. به طور کلی الگوریتم‌های طبقه‌بندی به دو دسته تقسیم‌بندی می‌شوند، الگوریتم‌های با نظارت و بدون نظارت. در الگوریتم‌های بانظارت در قسمت جمع‌آوری داده‌ها نیاز به استفاده از نمونه‌های آموزشی برای برآورد تخمین پارامتر اولیه طبقه‌بند و نمونه‌های آزمایشی برای برآورد کارایی این الگوریتم‌ها داریم در طرف مقابل در الگوریتم‌های بدون نظارت فقط نیاز به نمونه‌های آزمایشی می‌باشد و این الگوریتم‌ها توسط روش‌های خوشه‌بندی، کلاس‌ها را تفکیک می‌کنند.

الگوریتم‌های احتمال‌های پیشین کلاس‌ها $P(\omega_i)$ نیز اضافه شود به الگوریتم (MAP(maximum a posteriori probability)) تبدیل می‌گردد.

$$\arg \max g_i(x) = -\frac{1}{2}(x - \mu_i)' \Sigma_i^{-1} (x - \mu_i) - \frac{1}{2} \ln |\Sigma_i|$$

طبقه بندی سطوح موازی و KNN

این روش طبقه‌بندی بسیار ساده و سریع است که در اصل، با در نظر گرفتن هیستوگرام‌های مؤلفه‌های طیفی مجزا از نمونه‌های آموزشی، آموزش داده می‌شود. یکی از معایب این روش این است که ممکن است پیکسل‌های برخی نواحی طبقه بندی نشود یا برخی از آنها در چند کلاس همپوشانی داشته باشند این امر کمی پیچیدگی الگوریتم را افزایش می‌دهد [3]. در مقایسه، روش‌های حداقل فاصله و حداکثر احتمال تمام پیکسل‌های تصویر را طبقه‌بندی می‌کنند مگر اینکه از روش‌های آستانه‌ای استفاده شود. محدودیت دیگر این است که احتمال‌های پیشین کلاس‌ها استفاده نمی‌شوند، همان گونه که در روش حداقل فاصله نیز استفاده نمی‌شد. عیب دیگر اینکه، برای داده‌هایی با همبستگی زیاد، مکعب‌ها می‌توانند همپوشانی زیادی داشته باشند. از آنجایی که اضلاع مکعب‌ها موازی محورهای مختصات هستند، لذا بعضی مجموعه‌های داده‌ای وجود دارند که آنها را نمی‌توان تفکیک کرد. البته در تقسیم‌بندی الگوریتم‌های غیر پارامتری با نظارت در این مقاله از الگوریتم KNN(k-nearest neighbors) نیز استفاده خواهد شد. در این الگوریتم تعداد همسایه‌ها هم‌کلاس در یک فاصله مشخص معیار تصمیم‌گیری طبقه بند می‌باشد.

طبقه بندی حداقل فاصله

تابع ممیز طبقه‌بندی حداکثر احتمال را در حالت خاص با احتمال پیشین برابر برای تمامی کلاس‌ها در نظر بگیرید. طبقه‌بند حداکثر احتمال را می‌توان مانند طبقه‌بند حداقل فاصله سازمان دهی کرد، البته معیار فاصله‌ای که وابسته به جهت است و برای هر کلاس تغییر پیدا می‌کند. یکی از روش‌های طبقه بند حداقل فاصله که در این مقاله استفاده شده، طبقه‌بندی کننده

الگوریتم‌های استفاده شده به اختصار توضیح داده می‌شود.

الگوریتم‌های با نظارت

برای تحلیل کمی کامپیوتری اطلاعات تصویری سنجش از دور، معمولاً از روش طبقه بندی بانظارت استفاده می‌شود. این روشها متکی بر استفاده از الگوریتم‌های مناسبی هستند که برای تعیین برجسب عناصر تصویری برحسب نوع پوشش یا کلاس آنها استفاده می‌شوند. الگوریتم‌های متنوعی با روش‌های مختلف برای این کار طراحی گردیده‌است [5].

از الگوریتم‌هایی که مبتنی بر مدل‌های آماری و توزیع احتمال کلاس‌های مورد نظر هستند گرفته تا آنهایی که فضای چند طیفی را با استفاده از سطوحی که به صورت بهینه ارائه می‌دهند، تصویر را به زیر فضاهای مشخص از کلاس‌ها افراز می‌کنند. بدون در نظر گرفتن نوع الگوریتم، عمل طبقه‌بندی با نظارت شامل مراحل زیر است:

انتخاب مجموعه‌ای از پوشش‌های زمینی که باید در تصاویر مشخص شوند، که به آنها کلاس اطلاعاتی گفته می‌شود. انتخاب نقاط نماینده هر کلاس که نمونه آموزشی نیز نامیده می‌شوند. تخمین پارامترهای خاص طبقه‌بندی با استفاده از اطلاعات نمونه‌های آموزشی. تعیین پوشش زمینی (طبقه‌بندی) و برجسب زدن به تمام پیکسل‌های تصویر، با استفاده از سیستم آموزش داده شده. جمع‌بندی نتایج بصورت جدول یا نقشه‌های موضوعی. بنابراین تفاوت الگوریتم‌ها در انتخاب روش‌ها در مراحل ۳ و ۴ می‌باشد. در ادامه هر یک از روش‌های طبقه‌بندی بانظارت به اختصار بیان می‌گردد [5].

طبقه بندی با معیار حداکثر احتمال (MAP و ML)

این روش عمومی‌ترین روش طبقه‌بندی با نظارت است که در سیستم‌های سنجش از دور استفاده می‌شود. این روش را از لحاظ آماری می‌توان با شیوه‌های معیار تصمیم‌گیری حداکثر احتمال، مدل‌های گوسی چند متغیری سطوح تصمیم‌گیری و سطوح آستانه انجام داد [1]. همان طور که در رابطه (۱) مشاهده می‌شوند، این طبقه‌بند به تخمین دقیق بردار میانگین و ماتریس کوواریانس برای هر کلاس طیفی نیاز دارد. اگر به این

است. در این مقاله از دو الگوریتم شبکه عصبی با معماری مختلف استفاده شده است.

ماهالانویس با رابطه (۲) است.

$$\arg \min g_i(x) = (x - \mu_i)' \Sigma_i^{-1} (x - \mu_i)$$

طبقه‌بندی بدون نظارت (خوشه بندی)

خوشه‌بندی یک دسته‌بندی از پیکسل‌ها را در فضای چند طیفی ایجاد می‌کند. بنابراین پیکسل‌هایی که در یک خوشه خاص قرار می‌گیرند، از نظر طیفی شبیه به هم هستند. در طبقه بندی بدون نظارت یک تصویر به کلاس‌های نامعینی تقسیم بندی می‌شود. سپس این وظیفه استفاده کننده است که آن کلاس‌ها را برچسب گذاری کند. روش‌های زیادی برای خوشه‌بندی وجود دارند در این مقاله از دو روش خوشه‌بندی fuzzy c-mean و k-mean استفاده شده است.

مزیت ماهالانویس نسبت به حدکثر احتمالان است که ماهالانویس هم سریع‌تر است و هم مقداری از جهت را با استفاده از ماتریس کواریانس در نظر می‌گیرد که می‌تواند میانگین کلاس یا واریانس مشترک باشد. اگر به جای ماتریس کواریانس هر کلاس ماتریس کواریانس کل در رابطه لحاظ شود به روش طبقه‌بندی باهاتاچاریا (رابطه (۳)) خواهیم رسید.

$$\arg \min g_i(x) = \frac{1}{2}(x - \mu_i)' \Sigma^{-1} (x - \mu_i) + \frac{1}{2} \ln |\Sigma|$$

طبقه‌بندی با استفاده از شبکه‌های عصبی

شبکه‌های عصبی مصنوعی ساختارهای قدرتمندی برای فرایندهای کنترلی، طبقه‌بندی داده‌ها و خوشه‌بندی اطلاعات دارند. اوزان یک شبکه‌عصبی مصنوعی مرحله به مرحله بر اساس اطلاعات ورودی که دریافت می‌کند تغییر کرده و هر چه میزان این اطلاعات بیشتر شود شبکه خطای خود را بیشتر کاهش می‌دهد. یکی از مرسوم‌ترین انواع شبکه‌های عصبی، شبکه عصبی پرسپترون چندلایه است که به طور موفقیت آمیزی در بازه وسیعی از کاربردها از جمله طبقه‌بندی داده مورد استفاده قرار گرفته است. هنگام کار با شبکه‌های عصبی با دو مسئله روبه رو هستیم: انتخاب معماری مناسب و انتخاب الگوریتم آموزشی مناسب. معماری مناسب به معنی انتخاب بهینه تعداد لایه‌ها، تعداد نرون‌ها در هر لایه و نوع تابع تحریک هر نرون می‌باشد و معماری بهینه شبکه‌های عصبی مبتنی بر مجموعه داده‌ها و ویژگی‌های آنهاست. متداول‌ترین الگوریتم آموزشی این شبکه‌ها، الگوریتم پس انتشار خطا می‌باشد. در الگوریتم پس انتشار خطا در هر مرحله مقدار خروجی محاسبه شده جدید، با مقدار واقعی مقایسه شده و با توجه به خطای به دست آمده وزن‌های شبکه در جهت کم کردن خطای خروجی اصلاح می‌شود. اساس این کمینه سازی، حرکت بر روی بردار گرادیان تابع مربعات خطای شبکه می‌باشد. مشکل الگوریتم پس انتشار خطا، همگرایی کند و توقف در نقاط مینیموم محلی

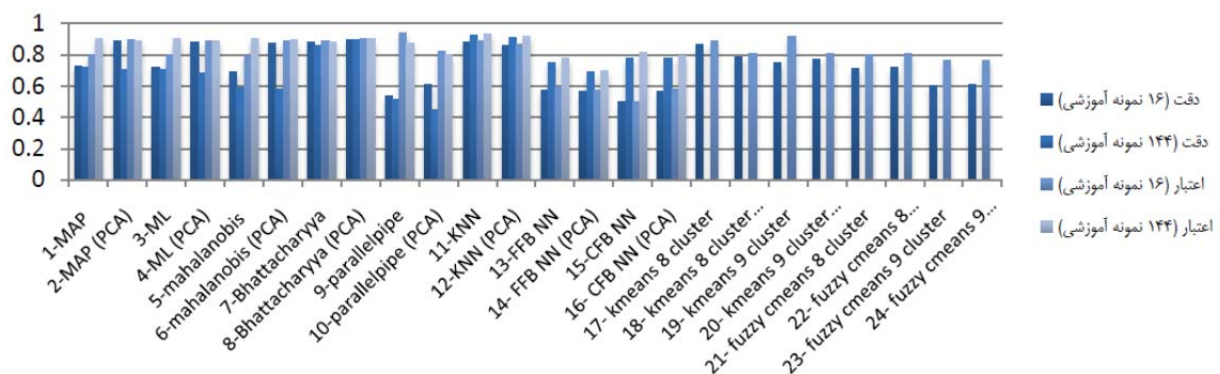
شبیه سازی

در این بخش هشت نوع طبقه‌بند با نظارت و دو نوع الگوریتم خوشه‌بندی که در بخش‌های قبل معرفی شدند با شبیه سازی با هم مقایسه خواهند شد. این شبیه سازی برای هر یک از هشت طبقه‌بند با نظارت با دو دسته نمونه آموزشی ۱۶ و ۱۴۴ عددی تست شده است. همچنین با استفاده از روش‌های کاهش ویژگی PCA و کاهش تعداد باندهای تصویر از ۱۲ به ۴ باند تمامی شبیه سازی‌ها تکرار شده است. البته برای هر کدام از الگوریتم‌های بدون نظارت خوشه‌بندی k-mean و fuzzy c-mean دو تست یک بار با ۸ خوشه و با دیگر با ۹ خوشه انجام گرفته است. در نهایت نتایج شبیه‌سازی‌های مختلف از نظر دقت، اعتبار و پیچیدگی در ۴۰ مورد مقایسه قرار می‌گیرند. نتایج شبیه سازی در جدول ۳ آمده است.

طبقه‌بندهای آماری (روش حداقل خطا MAP و روش حداکثر احتمال ML)، طبقه‌بند حداقل فاصله، طبقه‌بند سطوح موازی و طبقه‌بند نزدیک‌ترین k همسایه و طبقه‌بند شبکه عصبی و الگوریتم‌های بدون نظارت خوشه بندی k-mean و c-mean با یکدیگر مقایسه شده‌اند.

جدول ۳- نتایج شبیه‌سازی

	دقت (۱۶ نمونه آموزشی)	دقت (۱۴۴ نمونه آموزشی)	اعتبار (۱۶ نمونه آموزشی)	اعتبار (۱۴۴ نمونه آموزشی)	زمان (۱۶ نمونه آموزشی)	زمان (۱۴۴ نمونه آموزشی)
1-MAP	0.728	0.719	0.803	0.906	4.68	4.59
2-MAP (PCA)	0.893	0.705	0.896	0.893	2.49	2.42
3-ML	0.722	0.708	0.801	0.906	4.39	4.47
4-ML (PCA)	0.884	0.686	0.89	0.893	2.48	2.3
5-mahalanobis	0.691	0.59	0.797	0.905	3.37	3.28
6-mahalanobis (PCA)	0.874	0.579	0.888	0.897	1.76	1.7
7-Bhattacharyya	0.883	0.861	0.89	0.88	4.52	4.18
8-Bhattacharyya (PCA)	0.898	0.894	0.904	0.901	2.42	2.32
9-parallelpipe	0.537	0.516	0.942	0.875	0.192	0.175
10-parallelpipe (PCA)	0.612	0.447	0.826	0.802	0.186	0.186
11-KNN	0.881	0.925	0.893	0.93	4.92	39.65
12-KNN (PCA)	0.858	0.912	0.867	0.922	5.34	39.41
13-FFB NN	0.578	0.754	0.598	0.776	0.841	0.785
14- FFB NN (PCA)	0.566	0.692	0.575	0.699	0.09	0.093
15-CFB NN	0.498	0.78	0.499	0.818	0.112	0.0943
16- CFB NN (PCA)	0.57	0.783	0.58	0.796	0.095	0.091
17- kmeans 8 cluster	0.864		0.888		0.12	
18- kmeans 8 cluster (PCA)	0.786		0.809		0.048	
19- kmeans 9 cluster	0.751		0.921		0.2	
20- kmeans 9 cluster (PCA)	0.771		0.809		0.11	
21- fuzzy cmeans 8 cluster	0.716		0.8		1.26	
22- fuzzy cmeans 8 cluster (PCA)	0.722		0.809		0.57	
23- fuzzy cmeans 9 cluster	0.607		0.762		1.06	
24- fuzzy cmeans 9 cluster (PCA)	0.611		0.765		0.61	



شکل ۱- نمودار مقایسه الگوریتم‌های طبقه‌بند از نظر دقت و اعتبار

دلیل که کاهش ویژگی باعث کاهش تعداد طیف‌ها شده و ابعاد ماتریس‌های استفاده شده در محاسبات کوچک‌تر و در نتیجه سرعت پردازش بهبود می‌یابد.

در طبقه بندهای آماری و طبقه‌بندهای حداقل فاصله، استفاده از PCA با تعداد نمونه آموزشی کمتر، اعتبار را بهبود بخشیده این در حالی است که در طبقه بند نزدیک‌ترین همسایگی و سطوح موازی این مطلب صادق نیست به این دلیل که کاهش ویژگی با تعداد نمونه‌های کمتر، تصمیم‌گیری را برای طبقه‌بندی در این دو نوع طبقه بند دچار اختلال می‌کند. همچنین مشاهده می‌شود که استفاده از PCA در روش‌های خوشه‌بندی بهبودی در اعتبار میانگین بوجود نیاورده است. از نظر دقت نیز استفاده از PCA باعث بهبود در طبقه بندهای آماری شده ولی در روش‌های خوشه‌بندی بهبودی در دقت حاصل نشده است.

جمع بندی

روش حداقل فاصله نسبت به طبقه بندی حداکثر احتمال تکنیک سریع‌تری است اما به اندازه روش قبل قابل انعطاف نیست به این دلیل که از اطلاعات ماتریس کوواریانس استفاده نمی‌کند. تفاوت اصلی طبقه‌بند حداکثر احتمال و حداقل فاصله به استفاده از کوواریانس نمونه آموزشی بر می‌گردد. طبقه بند حداقل فاصله فقط بر مبنای فاصله نقطه از میانگین کلاس و بدون توجه به جهت پخش‌شان نسبت به میانگین نمونه را به یک کلاس برچسب می‌زند. طبقه بند حداکثر احتمال تصمیم خود را بر مبنای اطلاعات ماتریس کوواریانس تعدیل می‌کند. در نتیجه و بطور کلی از طبقه‌بند حداکثر احتمال کارایی بهتری مورد انتظار است. در عمل وقتی کوواریانس هر کلاس تحت تأثیر نویز سیستماتیک قرار می‌گیرد، روش حداکثر احتمال نتایج خوبی نخواهد داشت. در طبقه بند ماهالانویس مزیت آن نسبت به حداکثر احتمال آن است که هم سریع‌تر است و هم مقداری از جهت را با استفاده از ماتریس کوواریانس در نظر می‌گیرد که می‌تواند میانگین کلاس یا واریانس مشترک باشد. در کل می‌توان گفت روش سطوح موازی سریع‌ترین روش طبقه‌بندی است زیرا فقط مقایسه بین اجزای طیفی پیکسل با

این مقایسه در دو حالت با ۱۲ طیف و با استفاده از روش‌های کاهش ویژگی PCA و کاهش تعداد باندهای تصویر از ۱۲ به ۴ باند با دو دسته نمونه آموزشی ۱۶ به صورت تصادفی و ۱۴۴ عددی نمونه مجتمع (برای الگوریتم‌های با نظارت) و در مجموع ۴۰ مقایسه صورت گرفته است. این مقایسه بر روی دقت و اعتبار متوسط و مدت زمان پردازش برای هر طبقه بند انجام شده است. در شکل ۱ مقایسه نمودارهای مربوطه انجام شده و نتایج عددی متناظر در جدول ۳، آورده شده است.

همان طور که از نمودار فوق مشخص است طبقه بند شبکه عصبی دارای کمترین مقدار اعتبار میانگین در بین دیگر طبقه بندهاست. که این مطلب برای حالت انتخاب ۱۴۴ نمونه آموزشی کمتر از حالت ۱۶ نمونه آموزشی می‌باشد. به این دلیل که بالا بودن تعداد نمونه آموزشی باعث بهتر شدن عملکرد آموزش برای شبکه عصبی می‌باشد. طبقه‌بند سطوح موازی در این میان بیشترین مقدار اعتبار میانگین را به خود اختصاص داده است.

در ارتباط با دقت میانگین همان طور که از نمودار فوق مشخص است طبقه‌بند سطوح موازی دارای کمترین مقدار دقت میانگین بوده که این مطلب برای حالت انتخاب ۱۴۴ نمونه آموزشی مجتمع کمتر از حالت ۱۶ نمونه آموزشی تصادفی می‌باشد. چرا که تنها بیشتر کردن تعداد نمونه آموزشی به بهتر شدن دقت کمکی نمی‌کند بلکه نحوه انتخاب نمونه‌ها به گونه‌ای که تمامی ویژگی‌های آماری کلاس مورد نظر را بتواند توصیف کند تاثیرگذار است. طبقه‌بند نزدیک‌ترین همسایگی و بعد از آن طبقه بند حداقل فاصله بیشترین مقدار صحت میانگین را به خود اختصاص داده‌اند.

جدول ۳ مدت زمان پردازش بر حسب ثانیه را نیز برای همه مدل‌های طبقه‌بند نشان می‌دهد. مشخص است طبقه بندهای شبکه عصبی دارای کمترین مقدار زمان پردازش بوده و طبقه‌بند نزدیک‌ترین همسایگی بیشترین مقدار را دارد. البته برای طبقه بند نزدیک‌ترین همسایگی با کاهش تعداد نمونه‌ها با توجه به اینکه تعداد مقایسه در فرایند الگوریتم کمتر می‌شود، سرعت پردازش به شدت بهبود می‌یابد. استفاده از PCA جهت کاهش ویژگی باعث کاهش زمان پردازش می‌گردد، به این

2. Chapelle, O., Haffner, P., and Vapnik, V. N. (1999) Support vector machines for histogram-based image classification. *IEEE Transactions on Neural Networks*, 10, 1055-1064.
3. Huang, C., Davis, L. S., and Townshend, J. R. G. "An assessment of support vector machines for land cover classification." *International Journal of Remote Sensing*, (2002) 23, 725-749.
4. Tso, B. C. K., and Mather, P. M., (2001) *Classification Methods for Remotely sensed Data*. London: Taylor and Francis.
5. Wilkinson, G. G. (2000) *Processing and classification of satellite images*. Encyclopaedia of Analytical Chemistry, Edited by R. A. Meyers. John Wiley and sons, 8679-8693.
6. Zhu, G., and Blumberg, D. G., (2002) Classification using ASTER data and SVM algorithms; The case study of Beer Sheva, Israel. *Remote Sensing of Environment*, 80, 233-240.
7. N. Docquier, S. Candel, Combustion control and sensors: a review, *Progress in Energy and Combustion Science*, 28 (2002) 107-150

ابعاد طیفی box ها لازم است. از مقایسه نتایج در حالت ۱۴۴ نمونه آموزشی مجتمع با کاهش ویژگی به ۴ باند، نتایج بهتری حاصل شده است.

سازگاری مکانی در یک تصویر طبقه‌بندی شده را می‌توان با استفاده از این روش نزدیک‌ترین همسایگی توسعه داد. این نوع طبقه‌بند از طبقه‌بندهای با نظارت قبلی پیچیده‌تر است و این امکان را فراهم می‌آورد که ویژگی‌های مکانی یک ناحیه به طریقی منطقی در فرایند طبقه‌بندی بکار گرفته شود. این نوع طبقه‌بند گران‌ترین طبقه‌بند در دسته طبقه‌بندهای با نظارت است و نتایج حاصل از این پروژه از نتایج طبقه‌بند سطوح موازی بهتر شده است و این به قیمت افزایش زمان پردازش است. در حالت ۱۶ نمونه آموزشی با کاهش ویژگی نتایج بهتری به دست آمده است اما در ۱۴۴ نمونه آموزشی آزمایش شده کاهش ویژگی باعث بدتر شدن طبقه‌بندی گشته است. درخصوص بکارگیری شبکه‌های عصبی، اگرچه از نظر دقت و اعتبار نسبت به طبقه‌بند نزدیک‌ترین همسایگی کارایی پایین‌تری دارد، ولی باید توجه داشت که می‌توان این نوع طبقه‌بندها را با طبقه‌بندهای آماری در یک سطح در نظر گرفت علاوه بر این با توجه به معماری بکار گرفته شده در شبکه‌های عصبی می‌توان آنها را در کاربردهای آنلاین نیز بکار برد. اگر از زمان آموزش این نوع طبقه‌بندها صرف نظر کنیم، جزو سریع‌ترین طبقه‌بندها هستند. طبقه‌بند بدون نظارت k-mean زمان کمتری را در مقایسه با طبقه‌بند بدون نظارت c-mean برای پردازش نیاز دارد و سریع‌تر می‌باشد. و هم چنین این زمان برای حالتی که ۸ کلاستر انتخاب می‌کنیم بهتر از ۹ کلاستر است. مقدار دقت و اعتبار میانگین برای طبقه‌بند بدون نظارت k-mean در مقایسه با طبقه‌بند بدون نظارت c-mean بیشتر است. به طور کلی با توجه به اینکه از نمونه‌های آموزشی استفاده نمی‌کنند، نتایج قابل قبولی را ارائه می‌دهند.

مراجع

1. Abend, K., and Harley, T. J. (1969) Comments "On the mean accuracy of statistical pattern recognisers". *IEEE Transactions of Information Theory*, May, 420-421.

