



بهبود دقت طبقه‌بندی تصویر فراطیفی AVIRIS با استفاده از الگوریتم کاهش نویز رگرسیون بر مبنای پیش‌بینی فرضیه چندگانه

احسان لاله زاری^{۱*}، علی اسماعیلی^۲، سعید همایونی^۳

۱- دانشجوی کارشناسی ارشد مهندسی سنجش از دور، دانشگاه تحصیلات تکمیلی صنعتی و فناوری پیشرفته، کرمان

۲- استادیار گروه مهندسی سنجش از دور، دانشگاه تحصیلات تکمیلی صنعتی و فناوری پیشرفته، کرمان

۳- استادیار بخش جغرافیا، محیط‌زیست و ژئوماتیک دانشگاه اتاوا، کانادا

چکیده:

تصویربرداری فراطیفی، به عنوان یکی از فناوری‌های نوین سنجش از دوری، منبع ارزشمندی برای کاربردهای مختلف علوم زمین، از جمله تهیه نقشه‌های پوششی، شناسایی و اکتشاف معادن، نظارت زیست‌محیطی به شمار می‌رود. با این وجود، به دلایل سخت‌افزاری و فناوری این داده‌های تصاویری دارای مشکلات ذاتی هستند. از آنجایی که بهبود سیستم سخت‌افزاری سنجنده‌های فراطیفی بسیار پرهزینه است، روش‌های سنجش از دوری پردازش تصویر مانند کاهش نویز، استخراج ویژگی و غیره به دلیل هزینه کم و موثر بودن مورد توجه قرار گرفته‌اند. یکی از جدیدترین و کارآمدترین این روش‌ها، روش پیش‌بینی فرضیه چندگانه است. نقطه ضعف این روش عدم استفاده از روشی موثر در انتخاب باندهای با شباهت بیشتر است. هدف این مقاله بررسی روش پیش‌بینی فرضیه چندگانه و اتخاذ روشی مناسب برای انتخاب باندهای طیفی بر مبنای رگرسیون خطی است. به دلیل انعطاف زیاد روش پیش‌بینی فرضیه چندگانه در تعیین ضرایب شباهت بین باندها، برای انتخاب باندهای طیفی مشابه، این روش انتخاب و پیاده‌سازی شد. در این تحقیق صحت کلی طبقه‌بندی SVM برای مجموعه داده‌ی فراطیفی Indian Pines حاصل از روش پیشنهادی ۹۵/۸۲٪، بدست آمد که ۰/۴٪ بهبود را نشان داد و در طبقه‌بندی KNN صحت کلی ۹۲/۸۹٪ بدست آمد که ۸/۲۳٪ درصد افزایش داشت که نشان دهنده‌ی کارآمدی روش پیشنهادی است.

واژه‌های کلیدی: تصاویر فراطیفی، کاهش نویز، روش پیش‌بینی فرضیه چندگانه، SVM، KNN.



۱- مقدمه

تصویربرداری فراطیفی منبع ارزشمندی از اطلاعات مکانی برای گسترده‌ی فراوانی از کاربردها مانند، تهیه نقشه‌های پوششی، شناسایی و اکتشاف معادن، نظارت زیست‌محیطی است [۱]. با این وجود، تاثیرات نویز روی تصاویر فراطیفی باعث کاهش روش‌های استخراج اطلاعات و طبقه‌بندی این تصاویر می‌شود [۲]. یکی از مراحل مهم پیش پردازش تصاویر فراطیفی کاهش نویز و افزایش محتوای اطلاعاتی این تصاویر و کمک به بهبود قابلیت‌های تحلیلی تصاویر است [۳].

داده‌های فراطیفی به صورت ذاتی دارای میزان اطلاعات به نویز نسبتاً پایینی نسبت به داده‌های سنجش از دوری چندطیفی هستند [۴] به همین دلیل تا به امروز تحقیقات زیادی روی تاثیر کاهش نویز فراطیفی صورت گرفته است [۵-۷] که نمایان‌گر اهمیت انجام پیش‌پردازش کاهش نویز قبل از انجام پردازش‌هایی نظیر طبقه‌بندی، آشکارسازی هدف و غیره است [۸].

Bourenane و همکاران (۲۰۱۰) روشی بر مبنای بهبود الگوریتم فیلتر چندبعدی وینر (MWF) بر مبنای تنسور Tucker ارائه دادند، که به صورت همزمان روی مولفه‌های مکانی و طیفی تصاویر فراطیفی عمل می‌کند و به بهبود همزمان کیفیت تصویر و دقت طبقه‌بندی می‌پردازد. به طوری که باعث بهبود دقت طبقه‌بندی بوسیله‌ی ماشین بردار پشتیبان بر روی تصاویر فراطیفی سنجنده AVIRIS شد. با این حال استفاده از یک الگوریتم با هسته‌ی تنسور n حالتی، ممکن است به فشردگی اطلاعات و از دست دادن جزئیات مکانی منجر شود [۹].

Qian و همکاران (۲۰۱۲) روشی سازمان یافته برای مقابله با مشکل طبقه‌بندی تصاویر فراطیفی، ارائه کردند. روش آن‌ها شامل دو جزء کلی بود، یک توصیف کننده بافت مکانی-طیفی بر مبنای تبدیل موجک ۳ بعدی برای جمع‌آوری ویژگی‌های ذاتی و یک مدل رگرسیون منطقی پراکنده برای انتخاب ویژگی و طبقه‌بندی پیکسل‌ها. مزیت این روش ساده‌تر کردن تخمین پارامترهای آن و معتبر بودن آن روی داده‌های فراطیفی واقعی و کاهش نویز آن‌ها بود [۱۰].

Xu و همکاران (۲۰۱۳) نه تنها حذف همبستگی مکانی، بلکه حذف همبستگی طیفی را با استفاده از تبدیل موجک مد نظر داشتند. آن‌ها روشی بر مبنای رگرسیون چندخطی (MLR) ۲ و تبدیل موجک به منظور تخمین نویز تصاویر فراطیفی ارائه دادند که روی تعدادی از داده‌های شبیه‌سازی شده‌ی سنجنده AVIRIS و داده‌های واقعی سنجنده Hyperion اعمال شد. نتایج تجربی آن‌ها نشان داد که این روش انطباق و دقت بیشتری نسبت به سایر روش‌های طبقه‌بندی دارد [۱۱].

Chen و همکاران (۲۰۱۴) یک الگوریتم پیش‌پردازشی مقاوم به نویز تصاویر فراطیفی طراحی کردند که بر مبنای همسایگی پیکسل‌های مورد نظر عمل می‌کند. این روش می‌توانست دقت طبقه‌بندی هر دو طبقه‌بند بیشترین شباهت و ماشین بردار پشتیبان را بهبود دهد، مخصوصاً در شرایطی که اندازه نمونه‌ها کوچک انتخاب شود و تصاویر نیز نویز داشته باشند [۱۲].

تا کنون روش‌هایی که الگوریتم‌های کاهش نویز را به طور همزمان روی مکان و طیف تصاویر فراطیفی اعمال می‌کنند، دارای نتیجه‌ی بهتری روی دقت طبقه‌بندی این تصاویر بوده‌اند. بنابراین در تحقیق پیش رو، هدف ارائه

1-Multidimensional Wiener Filtering

2-Multi Linear Regression



الگوریتمی بهبودیافته به منظور کاهش نویز در تصاویر فراطیفی و ارزیابی الگوریتم مورد نظر در افزایش دقت طبقه‌بندی این تصاویر است.

۲- داده‌ها و روش‌ها

۲-۱- داده‌های مورد استفاده

داده فراطیفی Indian Pines توسط سنجنده AVIRIS از سایت آزمایشی شمال غربی ایالت ایندیانا جمع‌آوری شده است و شامل ۱۴۵ سطر و ۱۴۵ ستون در ۲۲۴ باند بازتاب طیفی در محدوده ۰.۴ تا ۲.۵ میکرومتر می‌باشد. واقعیت زمینی موجود در تصویر به شانزده کلاس تقسیم‌بندی شده (شکل ۱) و تعداد باندها به علت وجود ناحیه جذبی بخار آب به ۲۰۰ باند کاهش یافته است. باندهای کاهش یافته شامل باندهای ۱۰۴ تا ۱۰۸ و ۱۵۰ تا ۱۶۳ است [۱۳].



شکل ۱- نقشه‌ی واقعیت زمینی داده فراطیفی Indian Pines

به علت کافی نبودن تعداد پیکسل‌های بعضی از کلاس‌ها، تعدادی از آن‌ها حذف شده و به صورت بدون برچسب تلفیق شدند. در نهایت همانطور که در جدول ۱ این کلاس‌ها به صورت ضخیم و پر رنگ در آمده‌اند، ۹ کلاس باقی ماند.

جدول ۱ - تعداد پیکسل کلاس‌های داده فراطیفی Indian Pines

#	Class	Samples
1	Alfalfa	46
2	Corn-notill	1428
3	Corn-mintill	830
4	Corn	237
5	Grass-pasture	483
6	Grass-trees	730
7	Grass-pasture-mowed	28
8	Hay-windrowed	478
9	Oats	20
10	Soybean-notill	972
11	Soybean-mintill	2455
12	Soybean-clean	593
13	Wheat	205
14	Woods	1265
15	Buildings-Grass-Trees-Drives	386
16	Stone-Steel-Towers	93



۲-۲- روش شناسی

۱-۲-۲- پیش‌بینی فرضیه چندگانه (MH Prediction)

اگر تصویر فراطیفی را مجموعه‌ای از M بردار پیکسلی در نظر بگیریم: $X = \{x_m \in R^N, m = 1, 2, \dots, M\}$ به طوری که N تعداد باند‌های طیفی باشد. هدف، پیدا کردن یک ترکیب خطی بهینه برای تمام پیش‌بینی‌های ممکن یا فرض‌هایی برای نمایش مجدد x است. نمایش مجدد بهینه می‌تواند به صورت زیر فرموله شود:

$$\hat{W} = \operatorname{argmin} \|X - Hw\|_2^2 \quad (1)$$

به طوری که H یک ماتریس با ابعاد $N \times L$ است که ستون‌هایی با L فرض ممکن دارد و $\hat{W} = [\hat{w}_1, \dots, \hat{w}_L]^T$ یک بردار $L \times 1$ از ضرایب مربوط به تمام ستون‌های H است. به دلیل اینکه برای پنجره‌های با ابعاد بزرگ، $N < L$ است، ماهیت بد مطرح شده مسئله نیازمند بعضی از انواع منظم‌سازی است که برای تفکیک در میان تعداد بی‌شماری از ترکیبات خطی ممکن، تکیه بر فضای جواب معادله (۱) دارد. رایج‌ترین رویکرد برای منظم‌سازی مسئله کمترین مربعات، روش منظم‌سازی تیخونوف است [۱۴] که یک ضریب جریمه L_2 روی نرم \hat{W} اعمال می‌کند.

$$\hat{W} = \operatorname{argmin} \|X - Hw\|_2^2 + \lambda \|F_w\|_2^2 \quad (2)$$

به طوری که F ماتریس تیخونوف و λ پارامتر منظم‌سازی است. قسمت F اجازه تحمیل معلومات قبلی را روی جواب مسئله به ما می‌دهد. در این جا رویکردهایی پیشنهاد شده است که در آن فرض‌هایی که بیشترین عدم شباهت از بردار پیکسلی اصلی را دارند، نسبت به فرض‌هایی که بیشترین شباهت را دارند باید وزن کمتری به آن‌ها داده شود. به طور مشخص، یک F قطری به صورت زیر اتخاذ می‌شود:

$$F = \begin{bmatrix} \|x - h_1\|_2 & \dots & \emptyset \\ \vdots & \ddots & \vdots \\ \emptyset & \dots & \|x - h_L\|_2 \end{bmatrix} \quad (3)$$

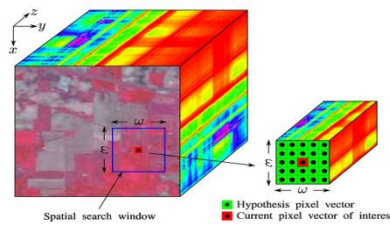
به طوری که h_1, h_2, \dots, h_L ستون‌های ماتریس H هستند. با این ساختار، F وزن‌های دامنه‌ی بزرگ اختصاص یافته به آن فرضیه‌هایی که دارای یک فاصله قابل توجه از x باشند را جریمه می‌کند. سپس برای هر بردار پیکسلی، \hat{W} می‌تواند مستقیماً بوسیله‌ی راه حل معمولی تیخونوف محاسبه شود:

$$\hat{W} = (H^T H + \lambda F^T F)^{-1} H^T x \quad (4)$$

بنابراین، بردار پیکسلی پیش‌بینی شده که x را تقریب می‌زند به صورت زیر محاسبه می‌شود:

$$\bar{x} = H \hat{W} \quad (5)$$

و مجموعه داده‌ی پیش‌بینی شده $\bar{X} = \{\bar{x}_m \in R^N, m = 1, 2, \dots, M\}$ بوسیله جایگذاری هر بردار پیکسلی در X با بردار پیش‌بینی شده مربوط به آن تولید می‌شود. برای یک بردار پیکسلی مورد نظر با در نظر گرفتن تمام همسایگی‌های بردار پیکسلی داخل یک پنجره، جستجوی مکانی با اندازه $w \times w$ تولید می‌شود (شکل ۵).



شکل ۲- پنجره‌ی جستجوی مکانی

سپس این بردارهای پیکسلی همسایگی، به عنوان ستون‌های ماتریس فرضیه H قرار داده می‌شوند. به دلیل اینکه باندهای طیفی تصویر فراطیفی دارای همبستگی هستند، می‌توانند به چندین گروه بر مبنای ضرایب همبستگی بین باندها تقسیم شوند، به طوری که در هر گروه با یکدیگر همبستگی بالایی داشته باشند [۱۲].

۲-۲-۲- آنالیز تفکیک‌پذیری فضای ویژگی (FSDA)

روش‌های استخراج ویژگی مانند LDA، GDA، NWFE^۵ و MMLDA^۶ تنها از دو اندازه‌گیری استفاده می‌کنند: پراکندگی درون کلاسی و پراکندگی بین کلاسی. ایمانی و قاسمیان (۲۰۱۵) الگوریتم آنالیز تفکیک‌پذیری فضای ویژگی (FSDA) را ارائه دادند [۱۵]. روش FSDA علاوه بر پراکندگی‌های درون کلاسی و بین کلاسی، از یک اندازه‌گیری سوم با عنوان پراکندگی بین طیفی نیز استفاده می‌کند. در تعداد نمونه‌های آموزشی کم روش FSDA دارای عملکرد بهتری نسبت به سایر روش‌ها است. در این روش بر خلاف سایر روش‌ها، به جای استفاده از میانگین کلاس‌ها از تمام نمونه‌های آموزشی برای تشکیل ماتریس‌های پراکندگی بین طیفی، بین کلاسی و درون کلاسی استفاده می‌شود که باعث افزایش کارایی طبقه‌بندی با تعداد نمونه‌های آموزشی کم می‌شود. در اغلب روش‌ها به دلیل محدودیت مرتبه ماتریس پراکندگی بین کلاسی ما میتوانیم حداکثر C-1 (تعداد کلاس‌ها است) ویژگی از داده‌ها استخراج کنیم ولی در روش FSDA به دلیل استفاده از نمونه‌های آموزشی به جای میانگین کلاس‌ها، این اجازه را به ما می‌دهد تا بیش از C-1 ویژگی از داده‌ها استخراج کنیم.

۳- نتایج

در این تحقیق سعی شده است که از رگرسیون خطی برای پیدا کردن باندهای مشابه استفاده شود. در انتخاب باند به روش رگرسیون، ابتدا نیاز به آماده‌سازی داده‌های باندهای طیفی است، به طوری که ابتدا ماتریس X از رابطه‌ی ۱۲ بدست می‌آید:

$$X = [x_1, B_1]; \quad (6)$$

به طوری که X ماتریس ضرایب آماده‌سازی شده، x_1 آرایه‌ای $1 \times N$ که N تعداد پیکسل‌های یک باند از تصویر است و B_1 باند شماره یک است که پیکسل‌های آن در یک ستون قرار گرفته‌اند. پس از تشکیل ماتریس X، ضرایب رگرسیون بین باند شماره یک (B_1) و باند شماره دو (B_2) از رابطه‌ی ۷ بدست می‌آیند:

$$A = (X^T X)^{-1} X^T B_2 \quad (7)$$

-
- 3-Linear Discriminant Analysis
 - 4-Global Discriminant Analysis
 - 5- Nonparametric weighted feature extraction
 - 6- Median-Mean Line based Discriminant Analysis



و سپس باند تخمین زده شده برای باند شماره دو (B_2) بوسیله ضرایب رگرسیونی از رابطه ۸ بدست می آید:

$$B' = A(1) + A(2) * B_1 \quad (8)$$

که $A(1)$ مولفه‌ی عرض از مبدا و $A(2)$ مولفه‌ی ضریب زاویه‌ی رابطه‌ی خطی می‌باشد. و در نهایت خطای این تخمین رگرسیونی از رابطه ۹ بدست می آید:

$$error = \|B_2 - B'\| \quad (9)$$

این الگوریتم دو به دو برای تمام باندها استفاده می‌شود و تصویری به ما می‌دهد که در آن باندهای با بیشترین شباهت، بیشترین ضریب رگرسیون را دارا می‌باشند. می‌توان از این تصویر برای دسته‌بندی باندهای طیفی بر مبنای شباهتشان استفاده کرد که نتایج آن در بخش بعدی خواهد آمد. ضروری است که در الگوریتم رگرسیونی خطای رابطه ۹ باید به حداقل برسد تا بهترین جواب را به ما بدهد. بنابر این این الگوریتم به صورت تکرار پذیر طراحی شده است تا به این هدف دست یابیم. پس از پیدا کردن باندهای با بیشترین شباهت طیفی از روش پیشنهادی، الگوریتم پیش‌بینی فرضیه چندگانه با استفاده از روش دسته‌بندی جدید توسعه پیدا کرد. مقایسه بین روش پیشنهادی و روش پیش‌بینی فرضیه چندگانه به منظور ارزیابی توانایی این روش‌ها در میزان کاهش نویز یا به عبارت دیگر افزایش نسبت سیگنال به نویز انجام گرفت. نسبت سیگنال به نویز روش‌های مختلف کاهش نویز از روش SNR به صورت زیر استفاده شد:

$$SNR(x_m, \hat{x}_m) = \log_{10} \frac{Var(x_m)}{MSE(x_m, \hat{x}_m)} \quad (10)$$

به طوریکه x_m بردار پیکسلی اصلی، \hat{x}_m بردار پیکسلی حاوی نویز و $Var(x_m)$ واریانس اجزای بردار x_m است. همچنین میانگین مربع خطا (MSE) به صورت

$$MSE(x_m, \hat{x}_m) = \frac{1}{M} \|x_m - \hat{x}_m\|_2^2 \quad (11)$$

بدست می‌آید [۱۲]. نتایج نسبت سیگنال به نویز در جدول ۳ مشاهده می‌شود.

جدول ۲- دسته‌بندی عددی باندهای طیفی دو مجموعه داده فراطیفی

روش پیش‌پردازش	مجموعه داده	دسته‌بندی عددی	نسبت سیگنال به نویز (SNR) دسی بل
MH	Indian Pines	{۱:۳۵ و ۳۶:۷۵ و ۷۶:۱۰۵ و ۱۰۶:۲۰۰}	۱۰/۸۰
Reg - MH ^y	Indian Pines	{۱:۱۴۵ و ۱۴۶:۲۰۰}	۲۴/۰۷

این بهبود نسبت سیگنال به نویز برای مجموعه داده‌ی Indian Pines ملموس‌تر است و این به طور کلی به خاطر پایین بودن نسبت سیگنال به نویز این مجموعه است. در مورد مجموعه داده‌ی Salinas نیز بهبود داشته ایم ولی به خاطر بالا بودن نسبت سیگنال به نویز مجموعه داده‌ی اصلی Salinas این بهبود به طور قابل توجهی چشمگیر نیست، اما همین مقدار بهبود نسبت سیگنال به نویز به نوبه‌ی خود باعث افزایش دقت طبقه‌بندی خواهد شد که در ادامه خواهیم دید. پس از پیش‌پردازش به روش پیشنهادی، الگوریتم استخراج ویژگی FSDA که قبل از این شرح آن داده شد روی این

⁷ - Regression Band Selection for Multi Hypothesis Prediction



مجموعه داده‌ی جدید به منظور بهبود تفکیک پذیری بین کلاسی اعمال شد. شرح مزایای الگوریتم استخراج ویژگی FSDA نسبت به سایر روش‌ها در بخش ۲-۲-۲ داده شده است.

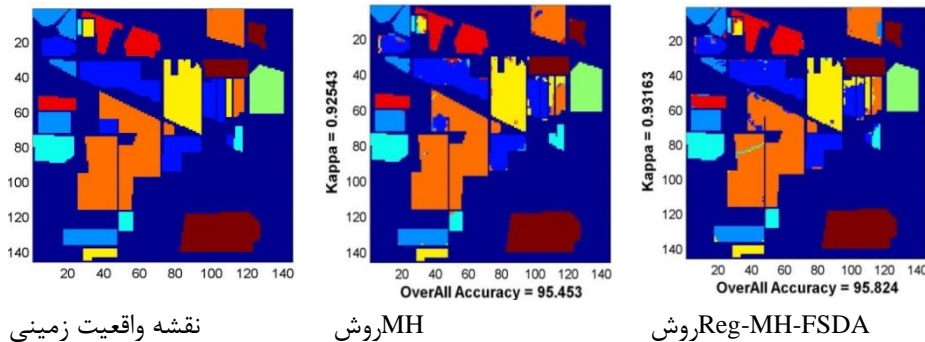
۳-۱- نتایج طبقه‌بندی به روش SVM

ابتدا به منظور اعتبارسنجی مدل برای جلوگیری از مسئله‌ی تطبیق بیش از حد و بدست آوردن پارامترهای بهینه‌ی طبقه‌بندی SVM با کرنل RBF، یعنی پارامترهای C و γ از طریق جستجوی شبکه‌ای، از اعتبارسنجی متقابل ۵ لایه استفاده شد که نتایج آن برای مجموعه داده‌های مختلف در جدول ۴ آمده است.

جدول ۳- پارامترهای بهینه‌ی طبقه‌بندی SVM برای مجموعه داده‌های فراطیفی

مجموعه داده	روش پیش پردازش	پارامترهای بهینه		دقت اعتبارسنجی
		C	γ	
Indian Pines	MH	۲۶۲۱۴۴	۰/۰۰۳۹	۹۹/۹۰
	Reg_MH	۱۰۲۴	۰/۱۲۵	۹۹/۹۱

کلاس‌های پیش‌بینی شده و نقشه‌ی واقعیت زمینی مجموعه داده‌ی Indian Pines در شکل ۴ آمده است.



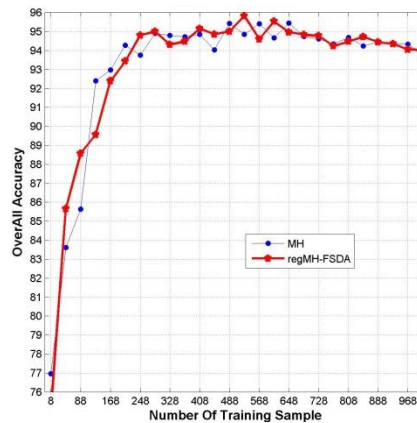
شکل ۴- کلاس‌های پیش‌بینی شده‌ی داده‌ی Indian Pines در مقابل واقعیت زمینی برای طبقه‌بندی SVM

صحت کلی و ضریب کاپا برای این طبقه‌بندی در جدول ۵ آمده است.

جدول ۴- صحت کلی و ضریب کاپای طبقه‌بندی SVM برای مجموعه داده‌ی Indian Pines

روش	صحت کلی	ضریب کاپا
MH	۹۵/۴۵	۰/۹۲۵۴
Reg-MH-FSDA	۹۵/۸۲	۰/۹۳۱۶

نمودار صحت کلی در برابر تعداد نمونه‌های آموزشی مختلف در شکل ۵ آمده است.



شکل ۵- نتایج طبقه‌بندی SVM برای تعداد نمونه‌های آموزشی مختلف

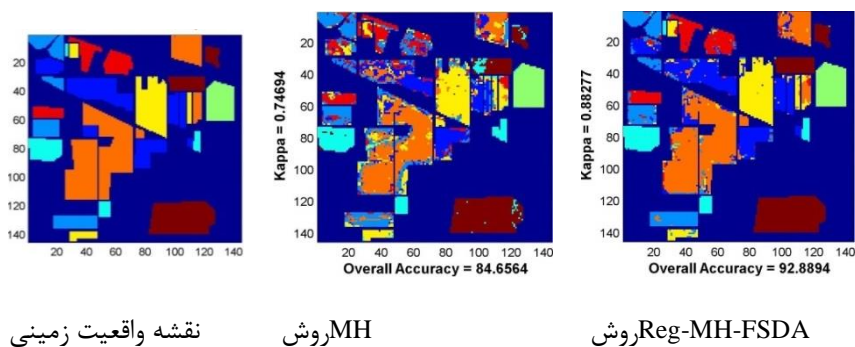
۳-۲- نتایج طبقه‌بندی به روش KNN

ابتدا از آنجایی که پس از تشکیل مدل طبقه‌بندی در روش KNN به ازای تعداد همسایگی‌های مختلف، دقت‌های طبقه‌بندی متفاوت می‌باشند از اعتبار سنجی متقابل ۵ لایه برای یافتن پارامتر بهینه‌ی تعداد همسایگی مجموعه داده‌ی فراطیفی با پیش‌پردازش‌های مختلف استفاده شد که نتایج آن برای مجموعه داده‌های مختلف در جدول ۷ آمده است.

جدول ۶- همسایگی‌های بهینه‌ی طبقه‌بندی KNN برای مجموعه داده‌های فراطیفی

مجموعه داده	روش پیش‌پردازش	همسایگی‌های بهینه	دقت اعتبارسنجی
Indian Pines	MH	۷	۹۹/۳۳
	Reg_MH	۱	۹۹/۸۵

کلاس‌های پیش‌بینی شده و نقشه‌ی واقعیت زمینی مجموعه داده Indian Pines برای طبقه‌بندی KNN در شکل ۶ آمده است.



شکل ۶- نتایج طبقه‌بندی KNN برای مجموعه داده‌ی Indian Pines

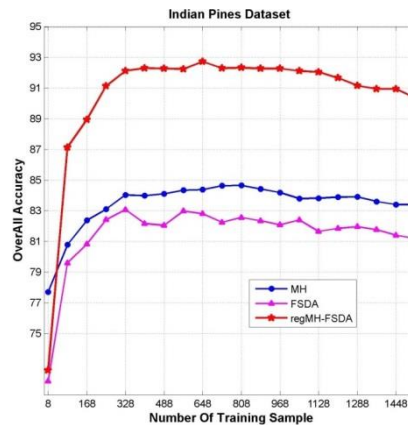
نتایج صحت کلی و ضریب کاپا برای طبقه‌بندی KNN این مجموعه داده در جدول ۷ قابل مشاهده می‌باشد.



جدول ۷- صحت کلی و ضریب کاپا برای طبقه‌بندی KNN

روش	صحت کلی	ضریب کاپا
MH	۸۴/۶۶	۰/۷۴۶۹
Reg-MH-FSDA	۹۲/۸۹	۰/۸۸۲۸

نمودار صحت کلی در برابر تعداد نمونه‌های آموزشی مختلف در شکل ۷ آمده است.



شکل ۷- نمودار صحت کلی در برابر تعداد نمونه‌های آموزشی مختلف در طبقه‌بندی KNN

۴- نتیجه‌گیری

روش پیش‌بینی فرضیه چندگانه با استفاده از امضاهای طیفی همسایگی پیکسل‌ها امضای طیفی پیکسل مورد نظر را پیش‌بینی کرده و این امر باعث افزایش چشمگیر نسبت سیگنال به نویز می‌شود. از طرفی دسته‌بندی بهینه‌ی باندهای طیفی مسئله‌ای بسیار تاثیرگذار در این الگوریتم است. در این تحقیق از روش پیش‌بینی رگرسیونی برای دسته‌بندی باندهای بهینه استفاده شده است که باعث بهبود ۱۳/۲۷ دسی بلی در نسبت سیگنال به نویز شده است که برای مجموعه داده‌ی Indian Pines که به طور کلی نسبت سیگنال به نویز پایینی دارد بسیار عالی عمل کرده است. این بهبود عملکرد در مورد دقت طبقه‌بندی SVM و KNN نیز مورد بررسی قرار گرفت. همانطور که از شکل ۵ پیداست الگوریتم پیشنهادی برای طبقه‌بندی SVM در تعداد نمونه‌های آموزشی مختلف به خوبی الگوریتم پیش‌بینی فرضیه چندگانه عمل کرده و حتی در بیشتر نقاط بهتر عمل کرده است. برای طبقه‌بندی KNN روش پیشنهادی بسیار عالی عمل کرده است (شکل ۷) به طوری که دقت طبقه‌بندی را ۸/۲۳ افزایش داده است.

این عملکرد خوب ناشی از تاثیرگذاری دسته‌بندی بهینه‌ی باندهای طیفی می‌باشد که حاصل استفاده از الگوریتم پیشنهادی است. بنابراین پیشنهاد می‌شود در پژوهش‌های آینده از این الگوریتم برای پیش‌پردازش سایر داده‌های فرایطی استفاده شود تا عملکرد آن روی داده‌های بیشتر بررسی شود.

مراجع

- [1] A. Plaza, J. A. Benediktsson, J. W. Boardman, and W. J. Blackwell, Recent advances in techniques for hyperspectral image processing, Remote Sens. Environ., vol. 113, no. S1, pp. 110–122, Sep. 2009.



- [2] . Skauli, Sensor noise informed representation of hyperspectral data with benefits for image storage and processing, *Opt. Exp.*, vol. 19, no. 14, pp. 13 031–13 046, Jul. 2011.
- [3] Zhang, H., Hyperspectral image denoising with cubic total variation model, *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* 7: 95-98, 2012.
- [4] O. Hisham, and S. Qian, Noise reduction of hyperspectral imagery using hybrid spatial-spectral derivative-domain wavelet shrinkage." *Geoscience and Remote Sensing, IEEE Transactions on* 44, no. 2 (2006): 397-408.
- [5] Chen, Guangyi, and Shen-En Qian, Simultaneous dimensionality reduction and denoising of hyperspectral imagery using bivariate wavelet shrinking and principal component analysis, *Canadian Journal of Remote Sensing* 34, no. 5 (2008): 447-454.
- [6] Chen, Guangyi, and Shen-En Qian, Denoising and dimensionality reduction of hyperspectral imagery using wavelet packets, neighbour shrinking and principal component analysis, *International Journal of Remote Sensing* 30, no. 18: 4889-4895, 2009.
- [7] Letexier, Damien, and Salah Bourennane, Noise removal from hyperspectral images by multidimensional filtering, *Geoscience and Remote Sensing, IEEE Transactions on* 46, no. 7: 2061-2069, 2008.
- [8] Phillips, Rhonda D., Christine E. Blinn, Layne T. Watson, and Randolph H. Wynne, An adaptive noise-filtering algorithm for AVIRIS data with implications for classification accuracy, *Geoscience and Remote Sensing, IEEE Transactions on* 47, no. 9: 3168-3179, 2009.
- [9] Bourennane, S., Fossati, C., & Cailly, A. (2010). Improvement of classification for hyperspectral images based on tensor modeling. *Geoscience and Remote Sensing Letters, IEEE*, 7(4), 801-805.
- [10] Qian, Y., Ye, M., & Zhou, J. (2013). Hyperspectral image classification based on structured sparse logistic regression and three-dimensional wavelet texture features. *Geoscience and Remote Sensing, IEEE Transactions on*, 51(4), 2276-2291.
- [11] Xu, D., Sun, L., & Luo, J. (2013). Noise estimation of hyperspectral remote sensing image based on multiple linear regression and wavelet transform. *Boletim de Ciências Geodésicas*, 19(4), 639-652.
- [12] Chen, C., Li, W., Tramel, E. W., Cui, M., Prasad, S., & Fowler, J. E. (2014). Spectral-spatial preprocessing using multihypothesis prediction for noise-robust hyperspectral image classification. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, 7(4), 1047-1059.
- [13] http://www.ehu.es/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes
- [14] Tikhonov, A. N., & Arsenin, V. I. (1977). *Solutions of ill-posed problems*. Vh Winston.
- [15] Imani, M., & Ghassemian, H. (2015). Feature space discriminant analysis for hyperspectral data feature reduction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 102, 1-13.



Hyperspectral Image Classification Accuracy Improvement Using Regression Noise Reduction Algorithm Based on Multi Hypothesis Prediction

Lalehzari, E. *¹, Esmaily, A. ², Homayouni, S.³

1- Ms.c student of remote sensing in Kerman University Of Advanced Technology.

2- Assistant professor in Department of Remote Sensing, Kerman University Of Advanced Technology.

3- Assistant professor in Department of Geography, Environment and Geomatics Engineering, University of Ottawa, Canada.

Abstract

Hyperspectral Imagery as a modern technology of remote sensing is a valuable source for various applications in geo-sciences; such as land cover maps, mine exploration and environmental monitoring. However, because of hardware and technological issues, these data have inherent problems. Since the improvement of hardware system in Hyperspectral sensors is very expensive, image processing methods like noise reduction and feature extraction, which are less expensive, are more important. The most recent and competent method is multi-hypothesis prediction. As its drawback, this method do not uses the suitable way to select the bands with maximum likelihood. The aim of this research is to apply the multi-hypothesis prediction method and to select the spectral bands based on linear regression. This approach was used because of its good flexibility in determining the similarity coefficient of spectral bands. The Support Vector Machine and K Nearest Neighbor classification methods were applied on hyperspectral data of Indian Pines. The results showed the overall accuracy of SVM Classification 95.82 and KNN Classification 92.89 for Proposed method in selected data set. This means an improvement of overall accuracy of 0.4 percent for SVM classification and 8.23 percent for KNN classification in data set. The results revealed the good performance of applied method.

Keywords: Hyperspectral Images, Noise Reduction, Multi-hypothesis Prediction, SVM, KNN, Classification

Correspondence Address: Remote Sensing Group, Department of Civil & Surveying, Kerman University of Advanced Technology, Kerman, Iran. **Tel: +98 3433776611-13.**

Email: ehsanlalehzari@gmail.com