

ارائه یک روش انتخاب ویژگی براساس الگوریتم ژنتیک و درخت تصمیم به منظور طبقه‌بندی تصاویر تمام پلاریمتریک راداری

ایمان خسروی^{۱*}، میرمجید موسوی^۲، جلال امینی^۳

- ۱- دانشجوی دکتری سنجش از دور، دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، پردیس دانشکده‌های فنی، دانشگاه تهران
۲- کارشناس ارشد سنجش از دور، دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، پردیس دانشکده‌های فنی، دانشگاه تهران
۳- دانشیار دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، پردیس دانشکده‌های فنی، دانشگاه تهران

تاریخ دریافت مقاله: ۹۴/۰۳/۲۶ تاریخ پذیرش مقاله: ۹۴/۱۰/۱۵

چکیده

یک تصویر تمام پلاریمتریک راداری (POLSAR) قادر است ویژگی‌های پلاریمتریک مهمی برای طبقه‌بندی پوشش زمینی فراهم کند. این ویژگی‌ها می‌توانند پارامترهای مستخرج از ماتریس پراکنش، کواریانس و هم‌دوسی یا پارامترهای مستخرج از روش‌های تجزیه هدف یا هر دو دسته باشد. در این مقاله، ویژگی‌های پلاریمتریک فراوانی از یک تصویر POLSAR استخراج می‌شود. سپس با استفاده از الگوریتم ژنتیک (GA) و درخت تصمیم (DT)، یک روش انتخاب ویژگی مبتنی بر طبقه‌بندی ارائه می‌شود. پس از آن، طبقه‌بندی‌کننده DT با ویژگی‌های انتخابی از روش پیشنهادی با طبقه‌بندی‌کننده DT با تمام ویژگی‌ها مقایسه می‌شود. علاوه بر این، روش پیشنهادی با روش انتخاب ویژگی GA و ماشین بردار پشتیبان (SVM) نیز مقایسه می‌شود. نتایج نشان داد که دقت روش پیشنهادی (DT) با ویژگی‌های منتخب از (GA-DT) حدوداً ۳ درصد بیشتر از دقت روش DT با تمام ویژگی‌ها و تقریباً نزدیک به دقت روش DT با ویژگی‌های منتخب از GA-SVM شد. این درحالی‌است، که سرعت عملکرد روش پیشنهادی تقریباً ۵ برابر بیشتر از سرعت عملکرد روش DT با ویژگی‌های منتخب از GA-SVM شد. به‌عنوان یک نتیجه‌ی دیگر، ویژگی‌های منتخب از روش پیشنهادی موفقیت بیشتری در تفکیک کلاس‌های شهری و پوشش گیاهی نسبت به ویژگی‌های دو روش دیگر داشتند.

کلیدواژه‌ها: انتخاب ویژگی، الگوریتم ژنتیک (GA)، درخت تصمیم (DT)، طبقه‌بندی، تصاویر تمام پلاریمتریک، رادار.

۱- مقدمه

از یک تصویر تمام پلاریمتریک راداری (*POLSAR*)، ویژگی‌های پلاریمتریکی فراوانی همچون ویژگی‌های ناشی از ماتریس‌های پراکنش^۱، همدوسی^۲ و کواریانس^۳ و همچنین ویژگی‌های ناشی از روش‌های تجزیه هدف^۴ در پایه پلاریزاسیون خطی و دایره‌ای قابل استخراج است [۱]. بنابراین به دلیل قابلیت تولید این حجم بالای اطلاعات، یکی از مهم‌ترین کاربردهای این تصاویر طبقه‌بندی پوشش زمینی است. تاکنون، الگوریتم‌های نظارت‌شده و نظارت‌نشده متعددی از جمله روش‌های پیشینه شباهت، فازی سی‌مینز، ویشارت، درخت تصمیم^۵ (*DT*)، شبکه عصبی^۶ (*NN*) و ماشین بردار پشتیبان^۷ (*SVM*) برای طبقه‌بندی این تصاویر ارائه شده است ([۲]-[۱۰] و [۱۲]-[۱۵]). هدف همه این مطالعات، نشان دادن قابلیت به‌کارگیری مؤثر ویژگی‌های پلاریمتریک فراوان مستخرج از تصاویر *POLSAR* در طبقه‌بندی پوشش زمینی بوده است.

با این حال، افزایش تعداد این ویژگی‌ها اگرچه می‌تواند منجر به تولید دقت بالایی در طبقه‌بندی شود، اما به‌طور هم‌زمان ممکن است باعث پیچیده‌تر شدن طبقه‌بندی و افزایش حجم محاسباتی آن، به‌خصوص در روش‌های پارامتریک شود. به‌گونه‌ای که در زمان ناکافی بودن تعداد نمونه‌های آموزشی نسبت به تعداد ویژگی‌ها، رخداد پدیده نفرین ابعاد ممکن خواهد شد [۱۱]. به همین منظور، در برخی مطالعات سعی شده تا از روش‌های انتخاب ویژگی جهت کاهش ابعاد فضای ویژگی این تصاویر استفاده شود. در مقاله مقصودی و همکاران (۲۰۱۳) یک روش انتخاب ویژگی با استفاده از یک معیار غیرپارامتریک مبتنی بر

محاسبه ماتریس‌های پراکنش بین‌کلاسی و درون‌کلاسی ارائه شد [۱۲]. در مقاله حدادی و همکاران (۲۰۱۱)، به کمک یک الگوریتم ژنتیک^۸ (*GA*) و طبقه‌بندی‌کننده *NN* و در مقاله صالحی و همکاران (۲۰۱۴) به کمک *GA* و *SVM* روشی برای انتخاب ویژگی از تصاویر رادارست-۲ ارائه شد ([۱۳]-[۱۴]). *NN* و *SVM* طبقه‌بندی‌کننده‌های کارایی هستند، اما *NN* دارای نتایج منحصره‌فردی نیست و با هر بار اجرای شبکه، به یک جواب غیرمشابه با جواب قبلی می‌رسد. *SVM* با وجود تولید نتایج منحصره‌فرد، اما به دلیل استفاده از هسته‌ها^۹ دارای زمان محاسباتی بالایی است. بنابراین بهتر است یک الگوریتم با سرعت بالاتر و کارایی مشابه با این روش‌ها در نظر گرفت. در مقاله کی و همکاران (۲۰۱۲)، الگوریتم *DT* برای انتخاب ویژگی‌های مفید در طبقه‌بندی شیء‌مبنای تصاویر *POLSAR* استفاده شد [۱۵]. این الگوریتم دارای سرعت بالاتری نسبت به *NN* و *SVM* است. اما ترکیب روش *GA* و *DT* در انتخاب ویژگی تصاویر *POLSAR* استفاده نشده است. ترکیب قوه ابتکاری الگوریتم *GA* و سرعت الگوریتم *DT* می‌تواند یک الگوریتم کارا و سریع برای انتخاب ویژگی از تصاویر *POLSAR* ابعاد بالا ارائه دهد.

بنابراین این پژوهش به دنبال آن است تا به جای استفاده از *NN* و *SVM* از *DT* به همراه *GA* در انتخاب ویژگی تصاویر *POLSAR* بهره بگیرد. همان‌طور که در پایان مقاله مشاهده خواهد شد، روش پیشنهادی از سرعت بالا و کارایی قابل قبولی در انتخاب ویژگی تصاویر *POLSAR* برخوردار است. ساختار ادامه مقاله بدین شرح است: بخش ۲ به‌طور مفصل به ویژگی‌های پلاریمتریک می‌پردازد. در این بخش، دو دسته ویژگی‌های مستخرج از داده اصلی و ویژگی‌های مستخرج از روش‌های تجزیه هدف معرفی

¹ Scattering² Coherency³ Covariance⁴ Target Decomposition⁵ Decision Tree⁶ Neural Network⁷ Support Vector Machine⁸ Genetic Algorithm⁹ Kernel

۲-۱. ویژگی‌های پلاریمتریک مستخرج از داده

اصلی

مهم‌ترین ویژگی‌ها و شاخص‌های مستخرج از داده اصلی یا به عبارتی از عناصر ماتریس پراکنش که در مقالات مورد استفاده قرار گرفته است، شدت پس-پراکنش دریافتی، نسبت بین عناصر ماتریس پراکنش، نسبت بین عناصر ماتریس پراکنش با کل توان پراکنش‌یافته و ضریب هم‌بستگی بین قطبش‌های پلاریمتریک است.

$$\begin{aligned} \sigma_{hh} &= 10 \log |S_{hh}|^2, & \sigma_{hv} &= 10 \log |S_{hv}|^2, & \sigma_{vv} &= 10 \log |S_{vv}|^2 & \text{رابطه (۱)} \\ \sigma_{rr} &= 10 \log |S_{rr}|^2, & \sigma_{rl} &= 10 \log |S_{rl}|^2, & \sigma_{ll} &= 10 \log |S_{ll}|^2 & \text{رابطه (۲)} \\ R_{hhvv} &= 10 \log \frac{|S_{hh}|^2}{|S_{vv}|^2}, & R_{hvhh} &= 10 \log \frac{|S_{hv}|^2}{|S_{hh}|^2}, & R_{hvvv} &= 10 \log \frac{|S_{hv}|^2}{|S_{vv}|^2} & \text{رابطه (۳)} \end{aligned}$$

(R_{hhhh}, R_{hvvv}) است. رابطه ۳، فرمول این ویژگی‌ها را نشان می‌دهد [۱۷]:

این ویژگی‌ها در تخمین کمیت‌های هواشناسی مانند تعیین اندازه قطرات باران و شدت بارش کاربرد فراوانی دارد [۱۸]. هم‌چنین آن‌ها می‌توانند اطلاعات سودمندی درباره پوشش زمینی فراهم کنند. به‌طور مثال در مطالعات مشاهده‌شده که R_{hhvv} شاخص مفیدی در تعیین میزان زبری سطوح زمینی است [۱۹]. هم‌چنین با توجه به تئوری ترکیب پلاریزاسیون، نسبت قطبش‌های پلاریمتریک در پایه دایره‌ای $(R_{rrll}, R_{rlrr}, R_{rlll})$ نیز قابل محاسبه است.

۲-۱-۳. نسبت بین عناصر ماتریس پراکنش با

کل توان پراکنش‌یافته

ویژگی‌های قابل تعریف دیگر با عناصر ماتریس پراکنش، نسبت هر قطبش پلاریمتریک به کل توان پراکنش‌یافته $(SPAN)$ است که از روابط ۵ و ۶ در دو پایه خطی (R_{hh}, R_{hv}, R_{vv}) و دایره‌ای (R_{rr}, R_{rl}, R_{ll}) قابل محاسبه است [۲۰]

می‌شود. سپس در بخش ۳، داده و منطقه مطالعاتی و روش پژوهش ارائه می‌شود. بخش ۴ به پیاده‌سازی روش‌ها و بررسی و مقایسه نتایج مربوط می‌پردازد. درنهایت، بخش ۵ به نتیجه‌گیری مقاله ختم می‌شود.

۲. ویژگی‌های پلاریمتریک

به‌طور کلی دو دسته ویژگی پلاریمتریک در این بخش مورد بحث قرار می‌گیرد. دسته اول ویژگی‌های پلاریمتریکی که مستقیم از داده اصلی یا به عبارتی از عناصر ماتریس پراکنش قابل استخراج هستند و دسته دوم ویژگی‌ها یا توصیف‌گرهایی که از تئوری‌های تجزیه هدف به دست می‌آید. این دو دسته ویژگی عبارتند از:

۲-۱-۱. شدت پس پراکنش دریافتی

سیستم‌های متداول تمام پلاریمتریک راداری، پس-پراکنش دریافتی را در چهار قطبش افقی-افقی (HH) ، قائم-قائم (VV) ، افقی-قائم (HV) و قائم-افقی (VH) ذخیره می‌کند. بنابراین ساده‌ترین ویژگی قابل استخراج از تصاویر پلاریمتریک، شدت پس پراکنش دریافتی است [۱۶]. هم‌چنین با داشتن ماتریس پراکنش در پایه خطی افقی و قائم (S_{hh}, S_{vv}, S_{hv}) و براساس تئوری ترکیب پلاریزاسیون^۱، می‌توان شدت پس پراکنش دریافتی را در هر پلاریزاسیون دلخواه از جمله پلاریزاسیون دایره‌ای (S_{rr}, S_{ll}, S_{rl}) محاسبه کرد. روابط ۱ و ۲، فرمول شدت پس پراکنش دریافتی در پایه خطی $(\sigma_{hh}, \sigma_{hv}, \sigma_{vh})$ و دایره‌ای $(\sigma_{rr}, \sigma_{rl}, \sigma_{lr})$ را نشان می‌دهد

۲-۱-۲. نسبت بین عناصر ماتریس پراکنش

ویژگی مهم اولیه دیگری که از تصاویر پلاریمتریک قابل استخراج است، نسبت قطبش‌های پلاریمتریک $(R_{hhvv},$

¹ Polarization Synthesis

$$R_{rrll} = 10 \log \frac{|S_{rr}|^2}{|S_{ll}|^2}, \quad R_{rtrr} = 10 \log \frac{|S_{rl}|^2}{|S_{rr}|^2}, \quad R_{rlll} = 10 \log \frac{|S_{rl}|^2}{|S_{ll}|^2} \quad \text{رابطه (۴)}$$

$$R_{hh} = \frac{|S_{hh}|^2}{SPAN}, \quad R_{hv} = \frac{|S_{hv}|^2}{SPAN}, \quad R_{vv} = \frac{|S_{vv}|^2}{SPAN} \quad \text{رابطه (۵)}$$

$$R_{rr} = \frac{|S_{rr}|^2}{SPAN}, \quad R_{rl} = \frac{|S_{rl}|^2}{SPAN}, \quad R_{ll} = \frac{|S_{ll}|^2}{SPAN} \quad \text{رابطه (۶)}$$

بنابراین ضریب همبستگی پلاریمتریک در پایه‌های خطی و دایره‌ای ویژگی‌های مناسبی برای تفکیک و طبقه‌بندی پوشش‌های گیاهی مختلف است [۷].

۲-۱-۵. ویژگی‌های ناشی از ماتریس مولر

هنگامی که از بردار استوکس به جای بردار پس‌پراکنش جهت توصیف پلاریزاسیون موج انتشاری و موج پس‌پراکنش‌یافته استفاده شود، آن‌گاه توان پس-پراکنش‌یافته توسط ماتریسی به نام ماتریس مولر (به جای ماتریس پس‌پراکنش) قابل تعریف است. بردار استوکس دارای چهار مؤلفه است و بنابراین ماتریس مولر که توسط این بردار ساخته می‌شود، یک ماتریس ۴×۴ (به جای ماتریس پس‌پراکنش ۳×۳) است. این ماتریس توسط رابطه ۸ قابل محاسبه است [۲۲].

$$M = A^* (S \otimes S^*) A^{-1}, \quad A = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & j & -j & 0 \end{pmatrix} \quad \text{رابطه (۸)}$$

که در آن رابطه، S ماتریس پراکنش است. بنابراین هر عنصر ماتریس استوکس می‌تواند از عناصر ماتریس پراکنش محاسبه شود که چهار عنصر مهم قطری آن توسط روابط ۹ محاسبه می‌شود [۲۲].

$$M_{11} = \frac{1}{4} \{ S_{hh} S_{hh}^* + S_{vv} S_{vv}^* + 2 S_{hv} S_{hv}^* \}$$

$$M_{22} = \frac{1}{4} \{ S_{hh} S_{hh}^* + S_{vv} S_{vv}^* - 2 S_{hv} S_{hv}^* \}$$

$$M_{33} = \frac{1}{2} S_{hv} S_{hv}^* + \frac{1}{2} \text{Re} \{ S_{hh} S_{vv}^* \}$$

$$M_{44} = \frac{1}{2} S_{hv} S_{hv}^* - \frac{1}{2} \text{Re} \{ S_{hh} S_{vv}^* \}$$

در مقاله آوکی و همکاران (۱۹۹۹) نشان داده شد که این ویژگی‌ها در تعیین مناطق شهری و ساختمان‌ها مناسب هستند [۲۰]. به‌عنوان مثال، مشاهده شد که در مناطقی که دارای ساختمان‌های چوبی با دیوارهای عمود بر خط دید رادار است، مقدار پارامتر R_{hh} بسیار بالاست. همچنین در مناطقی که ساختمان‌های بلند و باریک زیادی وجود دارد، مقدار پارامتر R_{vv} زیاد است.

۲-۱-۴. ضریب همبستگی پلاریمتریک

ضریب همبستگی یا همدوسی بین دو قطبش پلاریمتریک از رابطه

$$\rho_{XY,AB} = \frac{|\langle S_{XY}, S_{AB}^* \rangle|}{\sqrt{\langle S_{XY}, S_{XY}^* \rangle \langle S_{AB}, S_{AB}^* \rangle}} \quad \text{رابطه (۷)}$$

قابل محاسبه است که در آن، XY و AB نمایان‌گر پایه پلاریزاسیون قطبش‌هاست.

مقدار ضریب همبستگی در پوشش‌های زمینی مختلف، متفاوت است. همچنین شرط تقارن بازتاب، شاخص مهمی در تفکیک مناطق شهری از مناطق طبیعی محسوب می‌شود. زیرا طبق شرط تقارن بازتاب در مناطق طبیعی، رابطه $\langle S_{hh}, S_{hv}^* \rangle \approx \langle S_{hv}, S_{vv}^* \rangle \approx 0$ و

در مناطق شهری $\langle S_{hh}, S_{hv}^* \rangle \neq 0, \langle S_{hv}, S_{vv}^* \rangle \neq 0$ برقرار است [۲۱].

علاوه بر ضریب همبستگی پلاریمتریک در پایه خطی، پایه دایره‌ای آن نیز رفتار پوشش‌های زمینی مختلف را نشان می‌دهد. به‌عنوان مثال، مشاهده شده که در مناطق جنگلی رابطه $\langle S_{ll}, S_{rr}^* \rangle \approx 0$ برقرار است.

توجیه شده در ۴۵ درجه یا مکانیسم حجمی (شبیه پوشش گیاهی) است [۲۳]. البته پارامتر سوم یعنی توان پراکنش یافته از هدف‌های با مکانیسم پراکنش حجمی هم‌بستگی شدیدی با ویژگی شدت پس‌پراکنش دریافتی σ_{hv} دارد. بنابراین استفاده از آن اضافی است.

۲-۲-۲. ضرایب تجزیه کروگاگر

ضرایب حاصل از تجزیه کروگاگر، توان پراکنش یافته از هدف‌های کره‌ای، دوسطحی و مارپیچی^۱ مانند را بیان می‌کند که از رابطه ۱۱ قابل محاسبه است که در آن $|k_s|^2$ توان پراکنش یافته از مؤلفه کره‌ای مانند، $|k_d|^2$ توان پراکنش یافته از مؤلفه دو سطحی و $|k_h|^2$ توان پراکنش یافته از مؤلفه مارپیچی ماتریس پراکنش است [۲۴]. پارامتر اول یعنی توان پراکنش یافته از مؤلفه کره‌ای مانند یعنی $|k_s|^2$ در واقع همان شدت پس‌پراکنش دریافتی در پایه دایره‌ای σ_{rl} است. بنابراین استفاده از آن اضافی به نظر می‌رسد.

۳-۲-۲. ضرایب تجزیه فریمن

روش تجزیه غیرهمدوس فریمن، ماتریس کواریانس را به مجموع ماتریس‌های کواریانس مکانیسم‌های پراکنش تک‌سطحی، دوسطحی و حجمی تجزیه می‌کند [۲۵]. بنابراین می‌توان توان‌های پراکنش یافته از این نوع هدف‌ها را به‌عنوان ویژگی در نظر گرفت. در این جا نیز مؤلفه توان پراکنش یافته از هدف‌هایی با مکانیسم حجمی، هم‌بستگی شدیدی با شدت پس‌پراکنش دریافتی σ_{hv} دارد که از محاسبه آن صرف‌نظر می‌شود. توان پراکنش یافته از هدف‌هایی با مکانیسم تک‌سطحی و دوسطحی از رابطه ۱۲ قابل محاسبه است.

این چهار مؤلفه نیز به نوعی توان پس‌پراکنش یافته را نشان می‌دهد.

۲-۲. ویژگی‌های مستخرج از روش‌های تجزیه هدف

ویژگی‌های استخراجی از روش‌های تجزیه هدف می‌تواند اطلاعات مفیدی درباره مکانیسم‌های پراکنش فراهم کند. روش‌های تجزیه هدف به دو دسته همدوس^۱ و غیرهمدوس^۲ تقسیم‌بندی می‌شود. روش‌های تجزیه همدوس، ماتریس پراکنش را به مجموعه‌ای از ماتریس‌های پراکنش متناظر با اشیای ساده‌تر یا استاندارد تجزیه می‌کند. روش‌های تجزیه غیرهمدوس، ماتریس کواریانس یا همدوسی را به مجموعه‌ای از توصیف‌گرهای درجه دوم متناظر با اشیای ساده‌تر یا استاندارد تجزیه می‌کند. تجزیه پائولی و کروگاگر نمونه‌ای از دسته اول و تجزیه فریمن^۳، یاماگوچی^۴ و پارامترهای آنتروپی^۵، ناهمسانگردی^۶ و زاویه آلفای میانگین $(H/A/\alpha)$ نمونه‌ای از دسته دوم است.

۱-۲-۲. ضرایب تجزیه پائولی

ضرایب تجزیه پائولی بیان‌گر توان پراکنش یافته از هدف‌هایی با مکانیسم‌های پراکنش تک (فرد) سطحی^۷، دو (زوج) سطحی^۸ و حجمی^۹ است که از روابط ۱۰ قابل محاسبه است. که در آن $|a|^2$ توان پراکنش یافته از هدف‌هایی با مکانیسم تک (فرد) سطحی، $|\beta|^2$ توان پراکنش یافته از هدف‌هایی با مکانیسم دو (زوج) سطحی توجیه شده در صفر درجه و $|\gamma|^2$ توان پراکنش یافته از هدف‌هایی با مکانیسم دو سطحی

¹ Coherent

² Incoherent

³ Freeman

⁴ Yamaguchi

⁵ Entropy

⁶ Anisotropy

⁷ Single(Odd)-Bounce

⁸ Double(Even)-Bounce

⁹ Volume

¹⁰ Helix

$$|\alpha|^2 = \left| \frac{S_{hh} + S_{vv}}{\sqrt{2}} \right|^2, \quad |\beta|^2 = \left| \frac{S_{hh} - S_{vv}}{\sqrt{2}} \right|^2, \quad |\gamma|^2 = 2|S_{hv}|^2 \quad \text{رابطه (۹)}$$

$$|k_s|^2 = |S_{rl}|^2, \quad |k_d|^2 = \min(|S_{rr}|^2, |S_{ll}|^2), \quad |k_h|^2 = \text{abs}(|S_{rr}|^2 - |S_{ll}|^2) \quad \text{رابطه (۱۰)}$$

دوسطحی، حجمی و ماریپیچی تجزیه می‌کند. همچنین این روش شرط تقارن بازتاب را در نظر نمی‌گیرد [۲۶]. به‌همین منظور پارامترهای حاصل از آن به‌ویژه مؤلفه ماریپیچی که عمدتاً در مناطقی با ساختار پیچیده (همچون مناطق شهری) ظاهر می‌شود، به‌عنوان ویژگی‌هایی جهت شناسایی و تفکیک مناطق شهری از پوشش‌های گیاهی مناسب به نظر می‌رسد که از رابطه ۱۳ قابل محاسبه هستند که در آن f_d, f_s, f_c و f_v ضرایب یاماگوچی است که سهم مکانیسم‌های تک‌سطحی، دوسطحی، حجمی و ماریپیچی را در ماتریس کواریانس نهایی نشان می‌دهد.

$$P_s^Y = f_s(1 + |\beta|^2), \quad P_d^Y = f_d(1 + |\alpha|^2), \quad P_v^Y = f_v, \quad P_c^Y = f_c \quad \text{رابطه (۱۳)}$$

پارامترهای $H/A/a$ از تجزیه بردار ویژه ماتریس همدوسی به دست می‌آید که از رابطه

$$H = -\sum_{i=1}^3 p_i \log_3(p_i), \quad p_i = \frac{\lambda_i}{\sum_{k=1}^3 \lambda_k}, \quad A = \frac{\lambda_2 - \lambda_3}{\lambda_2 + \lambda_3}, \quad \bar{\alpha} = \sum_{i=1}^3 p_i \alpha_i \quad \text{رابطه (۱۱)}$$

به‌دست آمده از تجزیه بردار ویژه هستند.

۲-۲-۵-۳. ارتفاع پدستال^۱

ویژگی دیگری که از مقادیر ویژه ماتریس همدوسی به‌دست می‌آید، ارتفاع پدستال است. ارتفاع پدستال میزان کمترین انرژی در امضای پلاریمتریک (هم‌پلاریزاسیون^۲) است و از رابطه ۱۵ قابل محاسبه است. در واقع این پارامتر به نوعی میزان انرژی پلاریزه‌نشده در سیگنال بازگشتی را

$$P_s = f_s(1 + |\beta|^2), \quad P_d = f_d(1 + |\alpha|^2) \quad \text{رابطه (۱۲)}$$

که در آن f_d, f_s, α و β ضرایب فریمن است. گفتنی است که روش فریمن تنها برای مناطقی مناسب است که شرط تقارن بازتاب یعنی $\langle S_{hh} \cdot S_{hv}^* \rangle \approx \langle S_{hv} \cdot S_{vv}^* \rangle \approx 0$ در آن برقرار باشد. بنابراین، ویژگی‌های حاصل از روش فریمن برای تشخیص مناطق طبیعی مناسب و برای مناطق شهری و اشیای ساخت دست بشر نامناسب است [۲۱].

۲-۲-۴. ضرایب تجزیه یاماگوچی

روش یاماگوچی توسعه‌یافته روش فریمن است که ماتریس کواریانس را به ماتریس‌های کواریانس هدف‌هایی با مکانیسم‌های پراکنش تک‌سطحی،

۲-۲-۵. پارامترهای حاصل از تجزیه بردار ویژه

۲-۲-۵-۱. پارامترهای $(H/A/a)$

قابل محاسبه است، که در آن λ_i ها مقادیر ویژه ماتریس همدوسی است [۲۲]. علاوه بر این سه مؤلفه، در مقاله آقای لی و پوتیر (۲۰۰۹) ثابت شده که برخی ترکیب‌های H و A همچون $HA, H(1-A), A(1-H)$ و $(1-H)(1-A)$ می‌تواند به‌عنوان ویژگی‌هایی برای تشخیص و شناسایی مکانیسم پراکنش و پوشش‌های زمینی مفید باشد.

۲-۲-۵-۲. مقادیر ویژه ماتریس همدوسی

ویژگی‌های دیگر قابل استخراج از ماتریس همدوسی، مقادیر ویژه آن یعنی λ_1, λ_2 و λ_3 است که در واقع شاخص وزن هر یک از مکانیسم‌های پراکنش

^۱ Pedestal Height

^۲ Co-pol

ارائه یک روش انتخاب ویژگی براساس الگوریتم ژنتیک و ...
ایمان خسروی، میرمجید موسوی، جلال امینی

$$\psi = \frac{\min(\lambda_1, \lambda_2, \lambda_3)}{\lambda_1 + \lambda_2 + \lambda_3} \quad \text{رابطه (۱۵)}$$

جدول ۱، لیست تمام ویژگی‌های پلاریمتریک توضیحی در این بخش را نشان می‌دهد.

اندازه‌گیری می‌کند. اغلب پوشش‌های گیاهی از جمله جنگل‌ها، محیط غیرپلاریزه‌کننده محسوب می‌شوند و بنابراین مقدار ارتفاع پدستال در این نواحی بسیار زیاد است [۲۸].

جدول ۳: ویژگی‌های پلاریمتریک استخراج‌شده در این مقاله

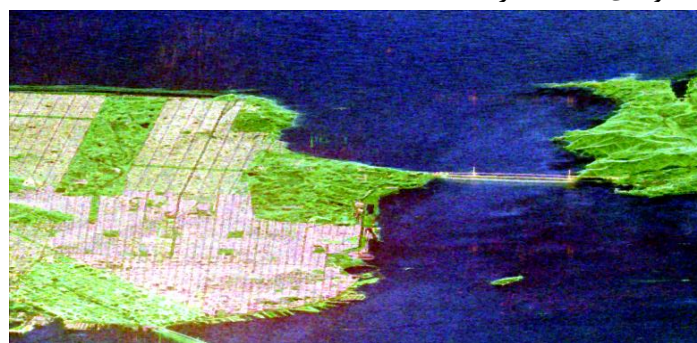
نام و علامت ویژگی‌ها				ویژگی‌ها
ضریب هم‌بستگی (۶)	نسبت بین عناصر با توان کل پراکنش (۶)	نسبت بین عناصر پراکنش (۶)	شدت پس‌پراکنش دریافتی (۶)	ویژگی‌های مستخرج از داده اصلی
$\rho_{hhhs}, \rho_{hvs}, \rho_{hsv}, \rho_{rvrl}, \rho_{rl}, \rho_{rll}$	$R_{hh}, R_{hv}, R_{vh}, R_{rv}, R_{rl}, R_{ll}$	$R_{hhhs}, R_{hvs}, R_{hsv}, R_{rvrl}, R_{rl}, R_{rll}$	$\sigma_{hh}, \sigma_{hv}, \sigma_{vh}, \sigma_{rv}, \sigma_{rl}, \sigma_{ll}$	عناصر قطری ماتریس مولر (۴)
			$M_{11}, M_{22}, M_{33}, M_{44}$	
ضرایب تجزیه یاماگوچی (۴)	ضرایب تجزیه فریمن (۲)	ضرایب تجزیه کروگاگر (۲)	ضرایب تجزیه پائولی (۲)	ویژگی‌های مستخرج از روش‌های تجزیه
$P_s^y, P_d^y, P_v^y, P_c^y$	P_s, P_d	$ k_d ^2, k_h ^2$	$ \alpha ^2, \beta ^2$	
	پارامترهای H/A/a (۷)	مقادیر ویژه ماتریس همدوسی (۳)	ارتفاع پدستال (۱)	
	$H, A, \bar{a}, HA, H(1-A), (1-H)A, (1-H)(1-A)$	$\lambda_1, \lambda_2, \lambda_3$	ψ	

اسمی آن در راستای رنج اریب ۱۱/۱ متر در رنج نزدیک و ۱۰/۵ متر در رنج دور است. تصویر رنگی پائولی تک دید منطقه مطالعاتی در شکل ۲ مشاهده می‌شود.

۳- مواد و روش

۳-۱- معرفی داده و منطقه مطالعاتی

داده استفاده‌شده در این مقاله، تصویر تمام پلاریمتریک راداری مربوط به باند C سنجنده ماهواره‌ای رادارست-۲ گرفته‌شده از سانفرانسیسکو بای است. قدرت تفکیک



شکل ۱. تصویر رنگی پائولی تک دید منطقه مورد مطالعه: سانفرانسیسکو بای، ماهواره رادارست-۲.

پس از انجام پیش‌پردازش، نمونه‌های آموزشی و آزمایشی هر کلاس از منطقه مطالعاتی انتخاب می‌شود. براساس داده مرجع منطقه، سه کلاس «مناطق شهری، پوشش گیاهی و آب» برای طبقه‌بندی این تصویر

قبل از طبقه‌بندی، بایستی اسپکل موجود در تصویر راداری کاهش یابد. به منظور حفظ لبه‌ها و خصوصیات پلاریمتریک، در این مقاله فیلتر لی روی ماتریس همدوسی تصویر موجود به کار می‌رود [۳۱].

۳-۲- روش پیشنهادی انتخاب ویژگی

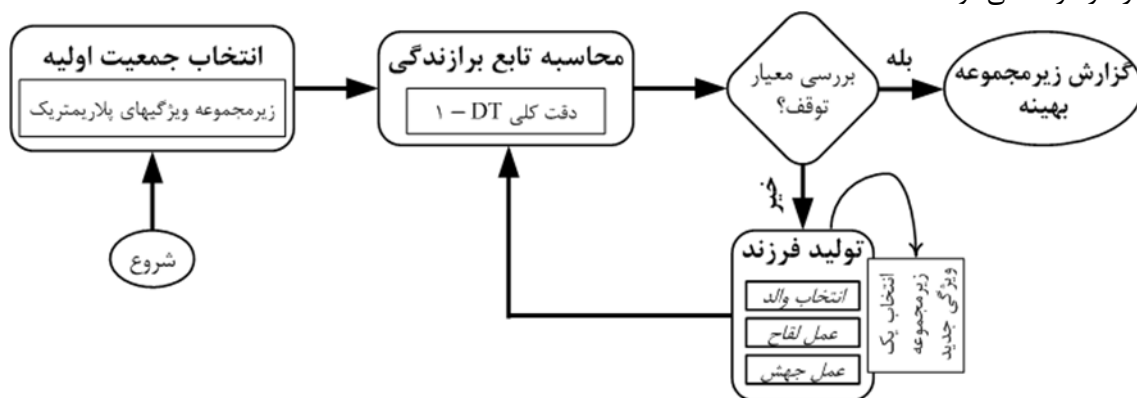
در این بخش، با استفاده از الگوریتم GA و DT ، اساس یک روش برای انتخاب ویژگی‌های پلاریمتریک مرتبط و حاوی اطلاعات در طبقه‌بندی ارائه می‌شود. در واقع، یک الگوریتم انتخاب ویژگی مبتنی بر طبقه‌بندی (روش رپر) ارائه می‌شود که در آن از الگوریتم GA برای ایجاد فضای جست‌وجو یا انتخاب زیرمجموعه ویژگی‌های داوطلب و از الگوریتم DT برای ارزیابی این زیرمجموعه‌ها استفاده می‌شود. چهارچوب روش پیشنهادی ارائه‌شده در شکل ۱ نشان داده شده است.

تعریف می‌شود. جدول ۲ نام هر کلاس و تعداد نمونه‌های آموزشی و آزمایشی آن را نشان می‌دهد.

جدول ۴: تعداد نمونه‌های آموزشی و آزمایشی کلاس‌های تعریف‌شده در منطقه مطالعاتی سانفرانسیسکو بای

کلاس	مناطق شهری	پوشش گیاهی	آب
# نمونه‌های آموزشی	۱۳۲	۱۴۹	۱۵۷
# نمونه‌های آزمایشی	۵۰۱	۵۷۶	۵۹۴

به‌منظور ارزیابی بهتر روش‌ها، تعداد نمونه‌های آزمایشی پنج برابر تعداد نمونه‌های آموزشی در نظر گرفته می‌شود.



شکل ۲: چهارچوب روش انتخاب ویژگی پیشنهادی در این مقاله

ارزیابی (مشهور به تابع برازندگی^۲ در ژنتیک)، جمعیت اولیه یا در واقع زیرمجموعه انتخاب‌شده اولیه، ارزش-گذاری می‌شود. اگر زیرمجموعه منتخب، معیار پیش-تعریف‌شده را برآورد کند، به‌عنوان زیرمجموعه ویژگی بهینه گزارش می‌شود. در غیر این صورت، با استفاده از دو عملگر مهم ژنتیک به نام عملگر لاقح^۳ و عملگر جهش^۴ فرزندان جدید یا همان زیرمجموعه ویژگی جدید انتخاب شده، مجدداً با به‌کارگیری تابع ارزیابی، ارزش‌گذاری آن زیرمجموعه انجام شده و در صورت برآورد کردن معیار، این زیرمجموعه گزارش می‌شود. در غیر این صورت، مراحل قبلی تا رسیدن به

الگوریتم GA از جمله الگوریتم‌های بهینه‌سازی کارا و پراستفاده در یادگیری ماشین و شناسایی الگو است. به کمک این الگوریتم، می‌توان به‌طور هم‌زمان به جست‌وجو در چندین نقطه از فضای تمام حالات ممکن ویژگی‌ها پرداخت. مطابق با شکل ۱، ابتدا در فضای جواب‌های ممکن برای ویژگی‌ها، تعدادی از زیرمجموعه‌های ویژگی به‌عنوان جمعیت^۱ اولیه انتخاب می‌شود. طول جمعیت برابر تعداد کل ویژگی‌های پلاریمتریک استفاده‌شده (در این جا ۴۹ ویژگی) است و به صورت باینری (۰ و ۱) وارد الگوریتم می‌شود؛ ۱ نشان‌دهنده انتخاب یک ویژگی و ۰ نشان‌دهنده عدم انتخاب آن ویژگی است. سپس، با اعمال یک تابع

² Fitness Function

³ Cross-over

⁴ Mutation

¹ Population

اطلاعاتی نمونه‌های S برای ویژگی A و $H(S)$ مقدار آنتروپی نمونه‌های S است. اگر ویژگی دارای c مقدار مختلف باشد، آن‌گاه آنتروپی S نسبت به این دسته‌بندی c گانه به صورت زیر تعریف می‌شود:

$$H(S) = -\sum_{i=1}^c p_i \log_2 p_i \quad \text{رابطه (۱۲)}$$

هم‌چنین به منظور جلوگیری از پیچیده‌تر شدن ساختار درخت از عمل هرس کردن درخت و کنار گذاشتن شاخه‌های اضافی و غیرضروری استفاده می‌کند [۳۰].

۴- پیاده‌سازی و ارزیابی نتایج

علاوه بر نمایش رنگی پائولی در شکل ۲، برخی ترکیب‌های دیگر ویژگی‌ها جهت نمایش رنگی منطقه مطالعاتی در شکل‌های ۳ مشاهده می‌شود. هم‌چنین ویژگی‌های انتخاب‌شده توسط روش پیشنهادی ($GA-DT$) در جدول ۳ نمایش داده شده است. برای ارزیابی روش پیشنهادی، آن را با طبقه‌بندی DT در حالت تمام ویژگی‌ها مقایسه می‌کنیم. هم‌چنین به منظور ارزیابی منصفانه‌تر، روش انتخاب ویژگی پیشنهادی با روش انتخاب ویژگی $GA-SVM$ مقایسه می‌شود. ویژگی‌های انتخاب‌شده این روش هم در جدول ۳ آورده شده است. برای مقایسه روش‌ها نیز از ماتریس خطا^۴، دقت کلی^۵ (OA)، دقت تولیدکننده^۶ (PA) کلاس‌ها و دقت کاربر^۷ (UA) کلاس‌ها (بر مبنای نمونه‌های آزمایشی) استفاده می‌شود. لازم به ذکر است که دقت PA هر کلاس به نوعی بیان‌گر درصد استخراج پیکسل‌های آن کلاس و دقت UA هر کلاس به نوعی بیان‌گر نرخ اعتمادپذیری پیکسل‌های استخراج‌شده آن کلاس است. نتایج ماتریس‌های خطا، دقت‌های OA ، PA و UA کلاس‌ها برای هر سه حالت در جدول‌های ۴ تا ۶ آورده شده است. هم‌چنین جهت مقایسه بصری دقت PA و UA

جواب نهایی ادامه می‌یابد [۲۹].

تابع برازندگی استفاده‌شده در روش پیشنهادی، مقدار دقت کلی یا مقدار خطای کل حاصل از طبقه‌بندی با الگوریتم DT روی نمونه‌های آزمایشی است که برای زیرمجموعه ویژگی بهینه انتخاب‌شده، مقدار اولی بایستی بیشینه و مقدار دوم بایستی کمینه شود. به عبارتی اگر F_i زیرمجموعه ویژگی مورد نظر باشد، تابع برازندگی یا ارزیابی آن به صورت

$$FF(F_i) = OA; \quad \text{or} \quad FF(F_i) = 1 - OA \quad \text{رابطه (۱۶)}$$

قابل تعریف است که در آن، FF تابع برازندگی و OA دقت کلی حاصل از طبقه‌بندی است.

در یادگیری ماشین نظارت‌شده، الگوریتم DT از درختان برای نمایش بصری تصمیم‌ها استفاده می‌کند. DT ، نمونه‌ها را به نحوی دسته‌بندی می‌کند که از گره ریشه^۱ به سمت پایین رشد کرده و در انتها به گره‌های برگ^۲ که همان تصمیم‌ها یا برچسب نمونه‌ها هستند، می‌رسد. هر گره داخلی یا غیربرگ، یک ویژگی را نشان می‌دهد که به تعداد ارزش (یا مقدار) آن ویژگی، شاخه از آن به گره‌های پایین‌تر وجود دارد. برای یادگیری این روش، الگوریتم‌هایی برپایه یک عمل جست‌وجوی حریصانه بالا به پایین وجود دارد که یکی از نمونه‌های تکامل‌یافته و کامل آن تحت عنوان $C5.0$ شناخته می‌شود. در این مقاله نیز از این الگوریتم برای یادگیری DT جهت طبقه‌بندی نمونه‌ها استفاده شده است. این الگوریتم از معیار بهره اطلاعاتی^۳ که مبتنی بر مفهوم آنتروپی است، جهت رشد درختان از بالا به پایین کمک می‌گیرد. این معیار از رابطه ۱۷

رابطه (۱۷)

$$Gain(S, A) = H(S) - \sum_{v \in \text{Values}(A)} \frac{|S_v|}{|S|} H(S_v)$$

قابل محاسبه است که در آن $Gain(S, A)$ بهره

⁴ Error Matrix

⁵ Overall Accuracy

⁶ Producer's Accuracy

⁷ User's Accuracy

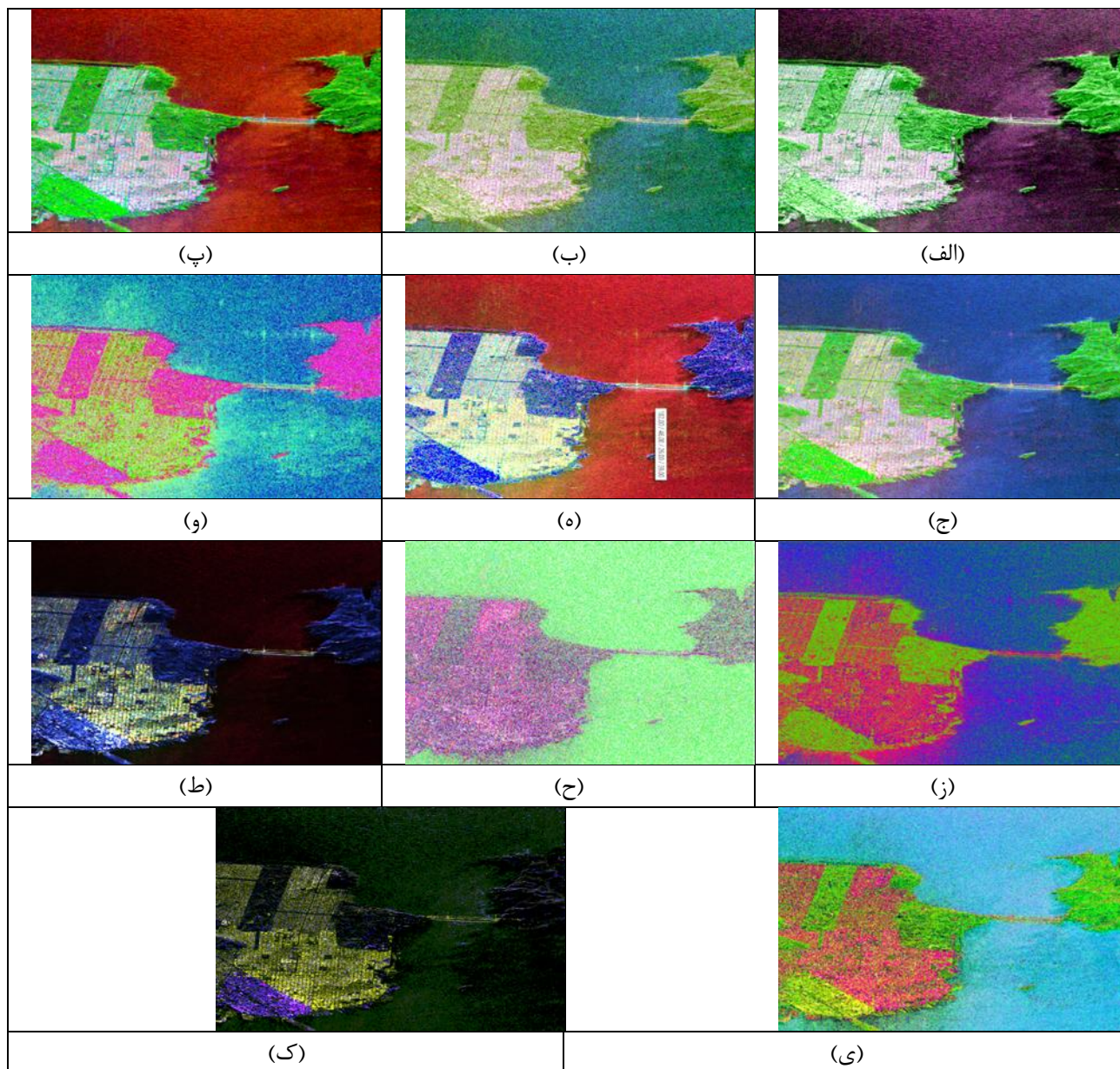
¹ Root Node

² Leaf Node

³ Information Gain

کلاس‌ها توسط هر سه حالت در شکل‌های ۴ الف و ۴ ب ارائه شده است.
 جدول ۵: ویژگی‌های انتخاب‌شده با استفاده از روش‌های مختلف انتخاب ویژگی

تعداد ویژگی‌ها	ویژگی‌های انتخاب‌شده	روش انتخاب ویژگی
۷	$\lambda_1, \lambda_3, \beta ^2, R_{hhvv}, R_{ll}, \rho_{rill}, \rho_{rrll}$	پیشنهادی: الگوریتم GA-DT
۸	$A, H, \beta ^2, R_{hv}, R_{rlrr}, \rho_{hhvv}, \sigma_{hh}, \sigma_{rr}$	الگوریتم SVM و GA

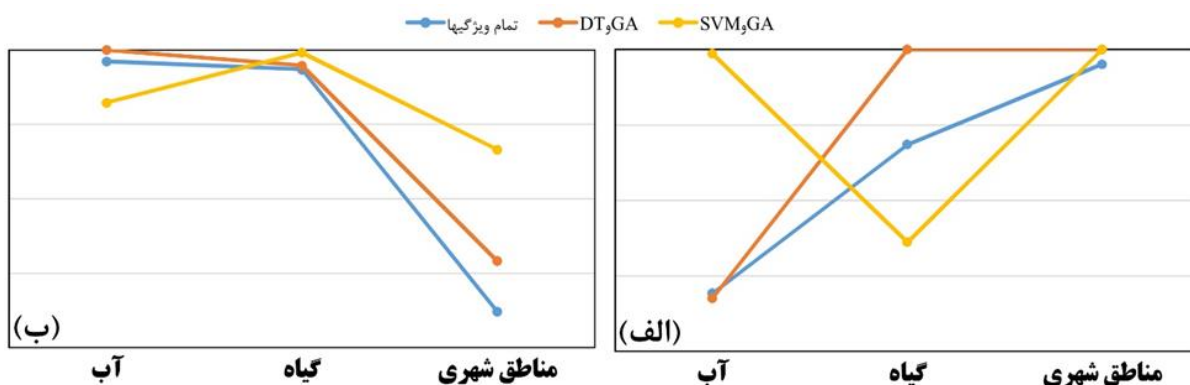


شکل ۳: نمایش‌های رنگی مختلف تصویر منطقه مطالعاتی. (الف) ترکیب رنگی سینکالر، (ب) ترکیب رنگی کروگاکر، (ج) ترکیب رنگی فریمن، (د) ترکیب رنگی یاماگوچی ۳، (ه) ترکیب رنگی یاماگوچی ۴، (و) ترکیب رنگی $H/A/a$ ، (ز) ترکیب رنگی ترکیبات چهارگانه H و A ، (ح) ترکیب رنگی ضرایب هم‌بستگی بین قطبش‌های پلاریزمتریک، (ط) ترکیب رنگی ۳ مقدار ویژه ماتریس هم‌دوسی، (ی) ترکیب رنگی IHS (شدت: SPAN، رنگ: α ، اشباع: $1-H$)، (ک) ترکیب رنگی عناصر قطری مولر ۴.

کسب کند. برای کلاس «آب» نیز تقریباً دقت تولیدکننده دو روش یکسان شده است (شکل ۴الف). بنابراین می‌توان گفت که به‌طور کلی روش پیشنهادی در استخراج پیکسل‌های کلاس‌ها نیز موفق‌تر بوده است. علاوه بر این، مطابق با این جداول و هم‌چنین با مشاهده شکل ۴ب، دقت کاربر هر سه کلاس حاصل از روش پیشنهادی خیلی بیشتر (۳ درصد تا ۷ درصد) از طبقه‌بندی DT با تمام ویژگی‌ها شده است که این مسئله بیان‌گر آن است که نتایج تولیدی از روش پیشنهادی در هر کلاس نیز قابل اعتمادتر از نتایج تولیدی از طبقه‌بندی DT با تمام ویژگی‌هاست. بنابراین براساس این دو نتیجه می‌توان بیان کرد که ویژگی‌های انتخاب‌شده، حاوی اطلاعات مفید و غیراضافی هستند که سبب تفکیک بهتر و قابل اعتمادتر کلاس‌ها از یکدیگر شده است.

همان‌طور که جداول ۴ تا ۶ نشان می‌دهد، دقت کلی حاصل از طبقه‌بندی DT با تمام ویژگی‌ها برابر با ۸۸/۰۳، طبقه‌بندی DT با ویژگی‌های انتخاب‌شده از روش پیشنهادی $GA-DT$ برابر با ۹۰/۷۸ و نهایتاً طبقه‌بندی DT با ویژگی‌های انتخابی از الگوریتم $GA-SVM$ برابر با ۹۳/۳۶ درصد شده است. طبق این نتایج، روش‌های انتخاب ویژگی توانسته‌اند دقت بهتری نسبت به تمام ویژگی‌ها به دست آورند. ضمن آن‌که سرعت عمل طبقه‌بندی در این دو حالت نیز نسبت به حالت اول به دلیل تعداد کم ویژگی‌ها بالاتر بود. (زمان محاسبات کمتر بود).

با مقایسه دقت‌های تولیدکننده و کاربر جدول ۴ و ۵ مشاهده می‌شود که برای کلاس‌های «شهری و گیاه»، طبقه‌بندی DT با ویژگی‌های انتخابی از $GA-DT$ ، توانسته دقت تولیدکننده بیشتری (۲ درصد تا ۸ درصد) نسبت به طبقه‌بندی DT با تمام ویژگی‌ها



شکل ۴: نمودار دقت‌های حاصل از چهار روش این مقاله، (الف) دقت تولیدکننده (PA)، (ب) دقت کاربر (UA)

کارایی کمتر در استخراج پیکسل‌های گیاه است. هم‌چنین دقت کاربر حاصل از روش DT با ویژگی‌های انتخابی از $GA-SVM$ برای کلاس «شهری» حدود ۲۲ درصد و دقت کاربر کلاس «گیاه» حدود ۳ درصد بیشتر از مقادیر نظیر آن‌ها در روش DT با تمام ویژگی‌ها شده است. هرچند که دقت کاربر کلاس «آب» حاصل از حالت اول حدود ۶ درصد بیشتر از حالت سوم است. بنابراین به‌طور ضمنی می‌توان بیان کرد که اعتمادپذیری نتایج منطقه شهری و گیاه حاصل از روش DT با ویژگی‌های انتخابی از $GA-SVM$ بیشتر از نتایج

با مقایسه جدول ۴ با ۶ مشاهده می‌شود که طبقه‌بندی DT با ویژگی‌های انتخابی از الگوریتم $GA-SVM$ ، توانسته دقت تولیدکننده خیلی بیشتری (۲ درصد تا ۲۰ درصد) نسبت به طبقه‌بندی DT با تمام ویژگی‌ها برای کلاس‌های «شهری و آب» کسب کند. اما برای کلاس «گیاه» دقت تولیدکننده آن حدود ۸ درصد کمتر از حالت اول شده است. بنابراین می‌توان گفت که روش DT با ویژگی‌های انتخابی از $GA-SVM$ دارای کارایی بیشتر نسبت به DT با تمام ویژگی‌ها در استخراج پیکسل‌های منطقه شهری و آب و دارای

مربوط از روش *DT* با تمام ویژگی‌هاست و برای آب این مسئله برعکس است.

جدول ۶: ماتریس خطا و دقت‌های ناشی از آن برای طبقه‌بندی *DT* با تمام ویژگی‌ها (۴۹ ویژگی)

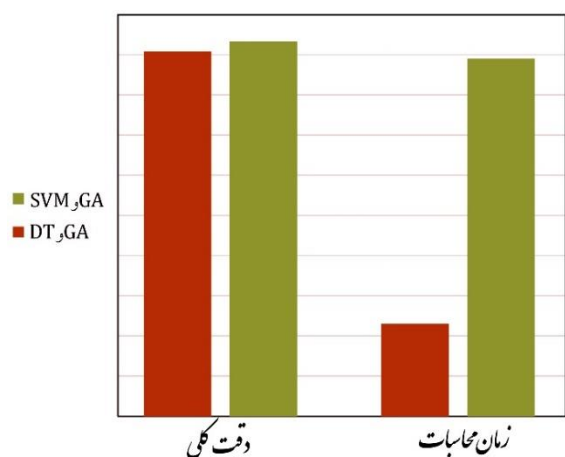
	مناطق شهری	گیاه	آب	UA
مناطق شهری	۳۲۵	۳۹	۱۳۷	۶۴/۸۷
گیاه	۴	۵۶۱	۱۱	۹۷/۴۰
آب	۰	۹	۵۸۵	۹۸/۴۸
PA	۹۸/۷۸	۹۲/۱۲	۷۹/۸۱	۸۸/۰۳

جدول ۷: ماتریس خطا و دقت‌های ناشی از آن برای طبقه‌بندی *DT* با ویژگی‌های انتخابی از پیشنهادی *GA-DT* (۷ ویژگی)

	مناطق شهری	گیاه	آب	UA
مناطق شهری	۳۵۹	۰	۱۴۲	۷۱/۶۶
گیاه	۰	۵۶۴	۱۲	۹۷/۹۲
آب	۰	۰	۵۹۴	۱۰۰/۰۰
PA	۱۰۰/۰۰	۱۰۰/۰۰	۷۹/۴۱	۹۰/۷۸

جدول ۸: ماتریس خطا و دقت‌های ناشی از آن برای طبقه‌بندی *DT* با ویژگی‌های انتخابی از *GA-SVM* (۸ ویژگی)

	مناطق شهری	گیاه	آب	UA
مناطق شهری	۴۳۴	۶۷	۰	۸۶/۶۳
گیاه	۰	۵۷۴	۲	۹۹/۶۵
آب	۰	۴۲	۵۵۲	۹۲/۹۳
PA	۱۰۰/۰۰	۸۴/۰۴	۹۹/۶۴	۹۳/۳۶



شکل ۵: مقایسه بین زمان محاسبات و دقت کلی الگوریتم *GA-DT* با الگوریتم *GA-SVM*

۵. نتیجه‌گیری

این مقاله همانند بسیاری از مقالات نشان داد که از یک

هم‌چنین مقایسه بین سرعت عملکرد و دقت کلی طبقه‌بندی *DT* با ویژگی‌های انتخابی از *GA-DT* با طبقه‌بندی *DT* با ویژگی‌های انتخابی از *GA-SVM* را نیز در شکل ۵ مشاهده می‌کنید. در این شکل به‌وضوح مشخص است که دقت کلی روش پیشنهادی این مقاله یعنی طبقه‌بندی *DT* با ویژگی‌های انتخابی از *GA-DT* تنها حدود ۲ درصد کمتر از دقت کلی طبقه‌بندی *DT* با ویژگی‌های انتخابی از *GA-SVM* است با این حال، سرعت عملکرد آن تقریباً حدود پنج برابر بیشتر از دومی است که این نکته می‌تواند به‌طور اساسی قابل توجه باشد. در واقع الگوریتم پیشنهادی این مقاله توانسته با سرعت خیلی بیشتر و دقت کلی تقریباً برابر نسبت به انتخاب ویژگی‌های حاوی اطلاعات مفید اقدام کند.

تفکیک مناطق «شهری و گیاه» و ویژگی‌های انتخاب شده توسط الگوریتم *GA-SVM* در تفکیک مناطق «شهری و آب» موفق‌تر از سایر مناطق عمل کرده‌اند. به‌علاوه، سرعت عملکرد روش پیشنهادی در این مقاله به‌مراتب (حدوداً ۵ برابر) بیشتر از سرعت عملکرد الگوریتم *GA-SVM* بود و دقت کلی و اکثر دقت‌های کاربری و تولیدی آن نیز نسبت به حالت «تمام ویژگی‌ها» خیلی بهتر بود.

تصویر تمام پلاریمتریک راداری، استخراج ویژگی‌های متعدد و مفیدی جهت طبقه‌بندی پوشش زمینی امکان‌پذیر است. در این مقاله، با استفاده از الگوریتم *GA-DT*، یک روش انتخاب ویژگی مبتنی بر طبقه‌بندی ارائه شد. علاوه بر این، به‌کمک الگوریتم *GA-SVM*، یک زیرمجموعه ویژگی دیگر انتخاب شد و نتایج حاصل از طبقه‌بندی با این ویژگی‌ها با حالت اول مقایسه گردید. طبق این نتایج می‌توان به‌طور کلی بیان کرد که ویژگی‌های انتخاب‌شده توسط الگوریتم *GA-DT* در

مراجع

- [1] C. Lardeux, P-L. Frison, C. Tison, J-C. Souyris, B. Stoll, B. Fruneau, and J-P. Rudant, "Support vector machine for multifrequency SAR polarimetric data classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 12, pp. 4143–4152, 2009.
- [2] J.A. Kong, A.A. Swartz, et al. "Identification of terrain cover using the optimal terrain classifier," *J. Electromagn. Waves Applicat.* vol. 2, pp. 171–194, 1988.
- [3] J.J. Van Zyl and C.F. Burnette, "Bayesian classification of polarimetric SAR images using adaptive a priori probability," *Int. J. Remote Sens.* vol. 13, no.5, pp. 835–840, 1992.
- [4] E. Rignot, R. Chellappa, and P. Dubois, "Unsupervised segmentation of polarimetric SAR data using the covariance matrix," *IEEE Tran. Geosci. Remote Sens.* vol. 30, no. 4, pp. 697–705, July 1992.
- [5] J.S. Lee, M.R. Grues, and R. Kwok, "Classification of multi-look polarimetric SAR imagery based on complex wishart distribution," *Int. J. Remote Sens.* vol. 15, pp. 2299–2311, 1994.
- [6] L. Ferro-Famil, E. Pottier, and J.S. Lee, "Unsupervised classification of multifrequency and fully polarimetric SAR images based on H/A/Alpha-wishart classifier," *IEEE Trans. Geosci. Remote Sens.* vol. 39, no. 11, pp. 2332–2342, November 2001.
- [7] T. Moriyama, S. Uratsuka, T. Umehara, M. Satake, A. Nadai, H. Maeno, K. Nakamura, and Y. Yamaguchi, "A study on extraction of urban areas from polarimetric synthetic aperture radar image," in *Proc. of IGARSS'04*, vol.1, pp.703–706, 2004.
- [8] C.J. Song, S. Yun, and L. Hui, "Classification of polarimetric SAR imagery based on target decomposition and neural network classifier," In *Proc. of ACRS '04*, November, 2004.
- [9] V. Alberga, "Comparison of polarimetric methods in image classification and SAR interferometry applications," Ph.D. dissertation, Tech. Univ. Chemnitz, Chemnitz, Germany, 2005.
- [10] C. Lardeux, P-L. Frison, J-P. Rudant, J.-C. Souyris, C. Tison, and B. Stoll, "Use of the SVM classification with polarimetric SAR data for land use cartography," in *Proc. of IGARSS'06*, July 31- August 04, Denver, Colorado, 2006.
- [11] I. Khosravi, and M. Mohammad-Beigi, "Multiple classifier systems for hyperspectral remote sensing data classification," *J. the Indian Society of Remote Sens.*, 42(2), pp. 423–428, 2014.
- [12] Y. Maghsoudi, M.J. Collins, and D.G. Leckie, "Radarsat-2 polarimetric SAR data for Boreal forest classification using SVM and a wrapper feature selector," *IEEE J. Selected Topics in Appl. Earth Observ. Remote Sens.*, vol. 6, no. 3, pp. 1531–1538, 2013.

- [13] A. Haddadi, M.R. Sahebi, and A. Mansourian, "Polarimetric SAR feature selection using a genetic algorithm," *Canadian J. Remote Sens.*, vol. 37, no. 1, pp. 27–36, 2011.
- [14] M. Salehi, M.R. Sahebi, and Y. Maghsoudi, "Improving the accuracy of urban land cover classification using Radarsat-2 POLSAR data," *IEEE J. Selected Topics in Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1394–1401, 2014.
- [15] Z. Qi, A. G. Yeh, X. Li, and Z. Lin, "A novel algorithm for land use and land cover classification using RADARSAT-2 polarimetric SAR data," *Remote Sens. Environ.*, vol. 118, pp. 21–39, 2012.
- [16] L. Pipia, V. Alberga, M. Chandra, M. Migliaccio, and K.B. Khadra, "Quantitative assessment of the efficiency of supervised classification using coherent and incoherent polarimetric SAR observables," In *Proc. of EUSAR 2002*, pp. 729–732, Cologne, Germany, June 2002.
- [17] V. Alberga and M. Chandra, "Analysis of amplitude ratios in SAR polarimetry," In *Kleinheubacher Berichte 2000*, pp. 527–534, Kleinheubacher, Germany, September 2000.
- [18] T.A. Seliga, V.N. Bringni, and H.H. Al-Khatib, "Differential reflectivity measurements in rain: first experiments," *IEEE Trans. Geosci. Remote Sens.* vol. 17, no. 4, pp. 240–244, October 1979.
- [19] J.C. Curlander, and R.N. McDonough, "Synthetic aperture radar, systems and signal processing," John Wiley and Sons, New York, USA, 1991.
- [20] H. Aoki, M. Matsuoka, and F. Yamazaki, "Backscattering characteristics of airborne SAR images for seismic vulnerability assessment in urban areas," In *Proc. of ACRS '99*, November, 1999.
- [21] S.V. Nghiem, S.H. Yueh, R. Kwok, and F.K. Li, "Symmetry properties in polarimetric remote sensing," *Radio Science*, vol. 27, no. 5, pp. 693–711, September 1992.
- [22] H.A. Zebker, J.J. vanZyl, "Imaging radar polarimetry: a review," *Proc. IEEE*, vol. 79, no. 11, pp. 1583–1606, November 1991.
- [23] S.R. Cloude and E. Pottier, "A review of target decomposition theorems in radar polarimetry," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, pp. 498–518, Mar. 1996.
- [24] E. Krogager, "A new decomposition of the radar target scattering matrix," *Electron. Lett.*, vol. 26, no. 18, pp. 1525–1526, 1990.
- [25] A. Freeman and S. Durden, "A three-component scattering model to describe polarimetric SAR data," in *Proceedings SPIE Conference on Radar Polarimetry*, Vol. 1748, pp. 213–225, San Diego, CA, July 1992.
- [26] Y. Yamaguchi, T. Moriyama, M. Ishido, and H. Yamada, "Four-component scattering model for polarimetric SAR image decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 8, pp. 1699–1706, August 2005.
- [27] J.S. Lee and E. Pottier, "Polarimetric radar imaging: from basics to applications," CRC Press, Taylor and Francis Group, New York, 2009.
- [28] K.M. Rodriguez, J.K. Weissel, and Y. Kim, "Classification of landslide surfaces using fully polarimetric SAR: examples from Taiwan," In *International Geoscience and Remote Sensing Symposium (IGARSS)* 5:2918–2920, 2002.
- [29] B. Bhanu, and Y. Lin, "Genetic algorithm based feature selection for target detection in SAR images," *Image and Vision Computing*, vol. 21(7), pp. 591–608, July 2003.
- [30] C. M., Bishop, "Pattern Recognition and Machine Learning," Springer, Second Edition, pp. 663–666, 2006.
- [31] J. S. Lee, M. R. Grunes, and G. de Grandi, "Polarimetric SAR speckle filtering and its implication for classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 5, pp. 2363–2373, 1999.



Presenting A Feature Selection Method Based On Genetic Algorithm and Decision Tree For Classifying Fully Polarimetric SAR Images

Iman Khosravi ^{*1}, Mir Majid Mousavi ², Jalal Amini ³

1- PhD. student of remote sensing in School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran

2- MA in remote sensing in School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran

3- Associate professor in School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran

Abstract

A fully polarimetric synthetic aperture radar (POLoSAR) image can provide important polarimetric features for land cover classification. These features can be the parameters obtained from scattering, covariance and coherency matrices, parameters extracted from target decomposition methods or both of them. In this paper, many polarimetric features are extracted from a POLoSAR image. Then, with the use of Genetic Algorithm (GA) and Decision Tree (DT), a feature selection method based on the classification is presented. Afterwards, a comparative analysis is accomplished between DT classification with features selected from the proposed method and DT classification with all features. Moreover, the proposed method should be compared with the feature selection method of GA and Support Vector Machine (SVM). The results indicated that the accuracy of the proposed method (DT classification with the features selected from GA-DT algorithm) is nearly 3% higher than the ones of the DT classification with all features and it is approximately equal with the ones of the DT classification with the features selected from GA-SVM algorithm. However, the performance speed of the proposed method is approximately 5 times more than the ones of DT classification with the features selected from GA-SVM algorithm. As an another result, the features selected from the proposed method have a more success than the ones of two other methods at classifying the urban areas and vegetation classes.

Key words: Regional gravity field modeling, Spherical Radial Basis Functions, Genetic algorithm, Tikhonov algorithm.