تفکر مکان محور برای پیشرفت جامعه

دومین همایش بین‌المللی تهیه نقشه و اطلاعات مکانی
و نوزدهمین همایش ملی ژئوماتیک

سازمان نقشه‌برداری کشور

محققین ارجمند:

مجتبی جنتی، میلاد نیرومندجدیدی، محمدجواد ولـــدان‌زوج، علی محمدزاده و علی‌رضا صفدری‌نژاد

بدین‌وسیله از حضور ارزنده شما و ارائه مقاله تحت عنوان: "استخراج نیمه‌اتوماتیک پارامترهای آماری کلاس‌های طبقه‌بندی مبتنی بر خوشه‌بندی تصویر" به صورت سخنرانی در **نوزدهمین همایش ملی ژئوماتیک ۹۱** تشکر و قدردانی می‌گردد.

ازدرگاه خداوند سبحان سلامتی، سعادت، و توفیق روز افزون برایتان آرزومندیم.

و من ا... التوفیق

هادی واعظی

معاون فنی سازمان نقشه برداری کشور

و

دبیر کل همایش

تذهیب Y 123 ، اثر محمدرضا هنرور، حق چاپ برای انتشارات یساولی محفوظ است.

استخراج نیمه‌اتوماتیک پارامترهای آماری کلاس‌های طبقه‌بندی مبتنی بر خوشه‌بندی تصویر

مجتبی جنتی[۱]، میلاد نیرومند جدیدی[۱]، محمدجواد ولدان‌زوج[۲]، علی محمدزاده[۳]، علی‌رضا صفدری‌نژاد[۴]

۱. دانشجوی کارشناسی‌ارشد مهندسی سنجش از دور، دانشکده مهندسی ژئوماتیک، دانشگاه صنعتی‌خواجه‌نصیرالدین‌طوسی
(m.jannati, niroumand)@sina.kntu.ac.ir
۲. دانشیار، گروه مهندسی سنجش از دور، دانشکده مهندسی ژئوماتیک، دانشگاه صنعتی‌خواجه‌نصیرالدین‌طوسی
valadanzouj@kntu.ac.ir
۳. استادیار، گروه مهندسی سنجش از دور، دانشکده مهندسی ژئوماتیک، دانشگاه صنعتی‌خواجه‌نصیرالدین‌طوسی
m_mohammadzadeh@kntu.ac.ir
۴. کارشناسی‌ارشد مهندسی سنجش از دور، دانشگاه صنعتی‌خواجه‌نصیرالدین‌طوسی
safdari_nezhad@sina.kntu.ac.ir

چکیده

اخذ نمونه‌های آموزشی مهمترین مرحله از فرآیند طبقه‌بندی نظارت شده محسوب می‌گردد. جمع‌آوری داده‌ها بر اساس مشاهدات زمینی و استفاده از داده‌های مرجع به عنوان روش‌های معتبر تهیه نمونه‌های آموزشی شناخته شده است. این امر معمولا مستلزم صرف وقت و هزینه بالا بوده و در برخی موارد نیز به دلیل در دسترس نبودن منطقه و یا استفاده از تصاویر آرشیوی، امکان‌پذیر نمی‌باشد. بدین ترتیب، تفسیر بصری عامل خبره به عنوان یک راهکار جهت اخذ بخشی از داده‌های آموزشی در نظر گرفته می‌شود. این روش مستعد عدم‌قطعیت‌های بالا در معرفی نمونه‌های آموزشی بوده و معمولا زمان‌بر می‌باشد. جهت رفع این مشکل، در تحقیق حاضر یک روش نیمه‌اتوماتیک برای استخراج پارامترهای آماری کلاس‌ها جهت استفاده در طبقه‌بندی نظارت شده ارائه گردیده است. بر اساس روش پیشنهادی، مشکل عدم‌ثبات پارامترهای آماری خوشه‌بندی k-means در یک فرآیند تکراری مرتفع شده و پارامترهای بهینه برای هر یک از کلاس‌ها برآورد گردیده است. پارامترهای برآورد شده به عنوان جایگزینی برای نمونه‌های آموزشی در نظر گرفته شده و دقت کلی طبقه‌بندی بر اساس پارامترهای برآورد شده در دو روش طبقه‌بندی بیشترین شباهت و ماشین‌های بردار پشتیبان به ترتیب ۹۱/۸۸ و ۹۲/۹۹ درصد و ضریب کاپا ۰/۸۷ و ۰/۸۹ برآورد گردید. همچنین یک سری نمونه آموزشی بر اساس تفسیر بصری عامل مجرب اخذ گردید و دقت طبقه‌بندی حاصل از آن با نتایج طبقه‌بندی به دست آمده از پارامترهای آماری برآورد شده، مقایسه شده است. به طور کلی پارامترهای آماری برآورد شده دقت قابل قبولی را در نتایج طبقه‌بندی نظارت شده به همراه داشته و می‌تواند به عنوان یک روش نیمه‌اتوماتیک برای استخراج پارامترهای آماری در نظر گرفته شود که تنها نیازمند معرفی تعداد کلاس‌های طبقه‌بندی می‌باشد.

واژه‌های کلیدی: طبقه‌بندی نظارت شده، نمونه‌های آموزشی، پارامترهای آماری، خوشه‌بندی

۱- مقدمه

به منظور تولید نقشه‌های پوششی/کاربری، روش‌های مختلف طبقه‌بندی تصاویر سنجش از دوری توسعه یافته است [۱- ۴]. در این روش‌ها یک مجموعه داده ناهمگن با ویژگی‌های متفاوت به چندین زیرمجموعه همگن که هر یک دارای ویژگی‌های مشابهی هستند، تفکیک می‌شود [۲]. این فرآیند براساس محاسبه یک معیار شباهت در فضای ویژگی[۱] n

بعدی داده‌ها صورت می‌پذیرد [۵]. اندازه‌گیری معیار شباهت برای هر یک از داده‌ها مستلزم در اختیار بودن یک مرجع مناسب می‌باشد. در طبقه‌بندی کننده‌های نظارت‌شده، این مرجع بر اساس پارامترهای آماری[۲] مستخرج از نمونه‌های آموزشی تعیین شده و میزان شباهت داده‌ها نسبت به آن سنجیده می‌شود. در طبقه‌بندی کننده‌های نظارت‌نشده، مرجع شباهت داده‌ها به صورت اتفاقی در فضای ویژگی انتخاب شده و در یک فرآیند تکراری تصحیح می‌گردد.

مهمترین مرحله از طبقه‌بندی نظارت‌شده، تعیین نمونه‌های آموزشی می‌باشد [۶]. به منظور جمع‌آوری داده‌های آموزشی، از مشاهدات زمینی [۷] و همچنین در صورت امکان از داده‌های مرجع مانند نقشه‌های توپوگرافی و تصاویر هوایی [۵]، استفاده می‌گردد. با توجه به هزینه بالای مشاهدات زمینی و همچنین در مواردی به دلیل در دسترس نبودن منطقه مورد نظر (صعب‌العبور بودن یا محدودیت مربوط به مرزهای سیاسی) و یا در صورت نیاز به استفاده از تصاویر آرشیوی فاقد نمونه‌های آموزشی (مانند تصاویر چندزمانه در استخراج تغییرات کاربری)، از تفسیر بصری عامل خبره برای استخراج نمونه‌های آموزشی استفاده می‌گردد [۵].

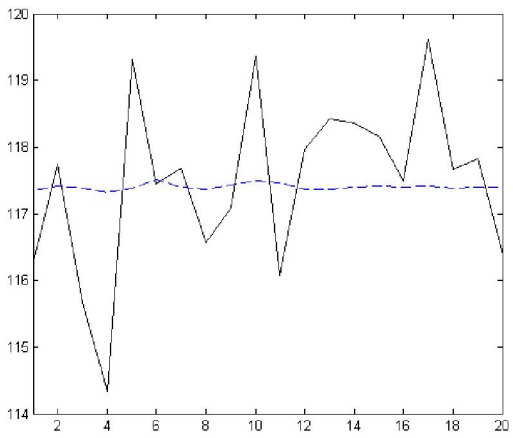
محدودیت‌هایی همچون زمان‌بر بودن و عدم امکان تفسیر همزمان بیش از سه بعد فضای ویژگی توسط عامل انسانی [۵] و امکان بروز اشتباه در شناسایی کلاس‌های با مشابهت طیفی [۶] در استخراج نمونه‌های آموزشی بر اساس تفسیر بصری موجب می‌گردد که این فرآیند دقت مطلوبی به همراه نداشته باشد. از این‌رو، در تحقیق حاضر یک روش نیمه‌اتوماتیک مبتنی بر خوشه‌بندی تصویر برای استخراج پارامترهای آماری کلاس‌ها پیشنهاد گردیده است. پارامترهای برآورد شده در فرآیند طبقه‌بندی نظارت‌شده مورد استفاده قرار گرفته و با نتایج حاصل از طبقه‌بندی با نمونه‌های آموزشی با کیفیت مناسب به دست آمده از تفسیر بصری مقایسه شده و همچنین با داده‌های چک زمینی مستقل مورد ارزیابی قرار گرفت.

مقاله پیش رو در چهار بخش کلی تدوین شده است. در بخش نخست مقدمه‌ای در خصوص اهمیت داده‌های آموزشی در فرآیند طبقه‌بندی نظارت‌شده و مشکلات مطرح در جمع‌آوری این داده‌ها ارائه گردید. در بخش بعد، روش پیشنهادی برای استخراج پارامترهای آماری کلاس‌ها تشریح شده است. در بخش سوم، نتایج حاصل از پیاده‌سازی ارائه گردیده است. در نهایت در بخش چهارم، دستآوردهای تحقیق مورد بحث و بررسی قرار گرفته است.

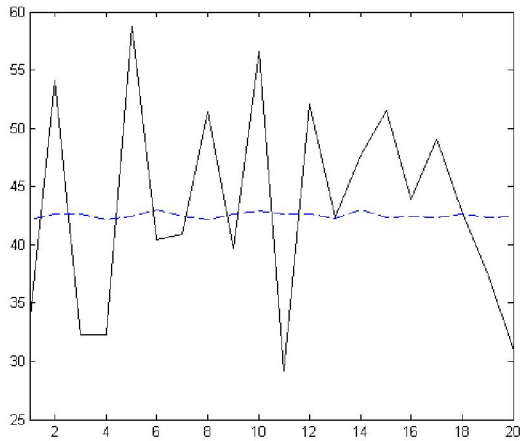
۲- متدولوژی

در تحقیق حاضر از خوشه‌بندی k-means برای استخراج پارامترهای آماری مورد استفاده در طبقه‌بندی نظارت‌شده بهره گرفته می‌شود. با توجه به تصادفی بودن مرکز اولیه خوشه‌ها در الگوریتم k-means، نتایج منحصر به فردی از این الگوریتم به دست نمی‌آید [۶ و ۸]. به منظور بررسی این موضوع، الگوریتم k-means به تعداد ۵۰۰ بار بر روی تصویر منطقه مطالعاتی پیاده‌سازی گردید. میانگین به دست آمده برای سه خوشه مورد نظر در باند ۵ تصویر در شکل (۱، خط پیوسته مشکی) نمایش داده شده است. بر این اساس مشاهده می‌شود که میانگین برآورد شده از الگوریتم k-means

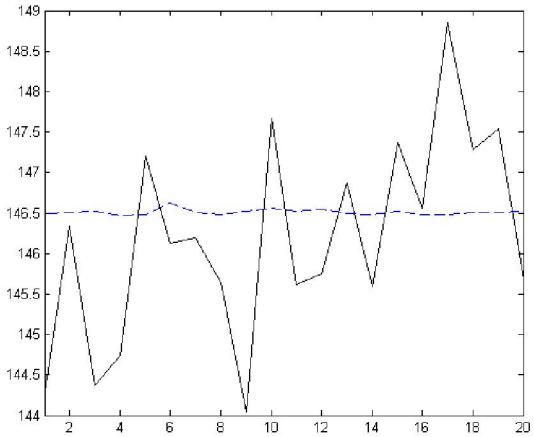
برای خوشه‌ها نه تنها منحصر به فرد نمی‌باشند، بلکه با نوسانات شدیدی نیز همراه هستند. به منظور حل این مشکل، الگوریتم k-means در یک فرآیند تکراری بر روی مجموعه‌های تصادفی ۱۰۰۰ عضوی از پیکسل‌های تصویر پیاده‌سازی گردید. این فرآیند تا زمانی که میانگین پارامترهای خوشه‌ها در تکرارهای مختلف به پایداری (تغییرات اندک) برسند، ادامه می‌یابد. میانگین پارامترهای آماری (میانگین و انحراف معیار) به ازای تکرارهای مختلف الگوریتم به عنوان پارامترهای بهینه در نظر گرفته شد. در نهایت به منظور ارزیابی ثبات این پارامترها، فرآیند انتخاب پارامترهای بهینه نیز ۲۰ بار تکرار شده و اثبات گردید که با پیاده‌سازی مستقل الگوریتم پیشنهادی، پارامترهای هر خوشه به صورت پایدار برآورد می‌گردد. مطابق شکل (۱، خط‌چین آبی) مشاهده می‌شود که میانگین بهینه خوشه‌ها در تکرارهای مختلف تغییرات چندانی نداشته و تقریبا پایدار می‌باشد. همچنین پارامتر انحراف معیار خوشه‌ها از ثبات قابل قبولی در تکرارهای مختلف برخوردار هستند. شایان ذکر است که این شرایط در تمام باندهای تصویر برقرار می‌باشد.



ب: پارامترهای مستخرج برای خوشه دوم



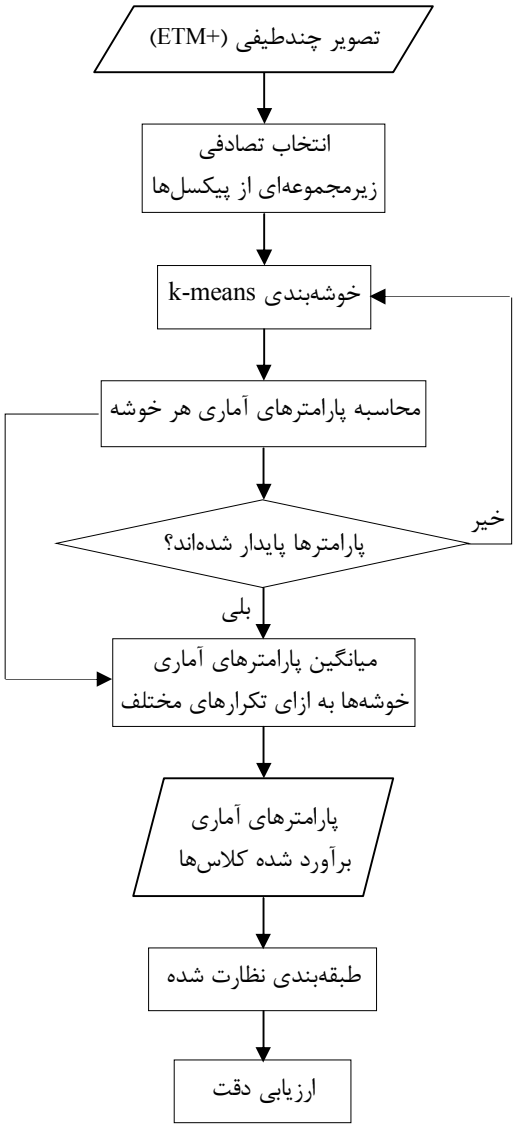
الف: پارامترهای مستخرج برای خوشه اول



ج: پارامترهای مستخرج برای خوشه سوم

شکل ۱: ارزیابی میزان پایداری میانگین برآوردشده توسط الگوریتم k-means (خط پیوسته) و میانگین بهینه در تکرارهای مختلف (خط‌چین)

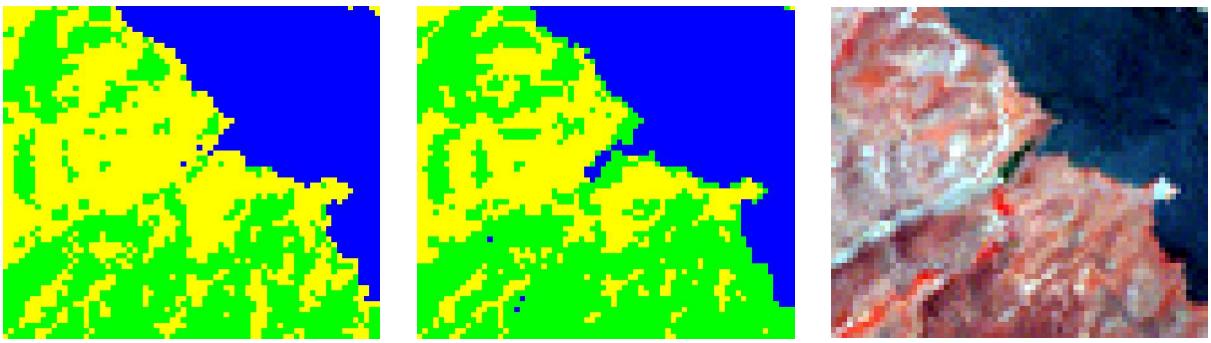
بدین ترتیب از میانگین پارامترهای آماری خوشه‌ها در تکرارهای مختلف الگوریتم k-means، به عنوان پارامترهای بهینه در طبقه‌بندی نظارت‌شده استفاده گردید (شکل ۲). به منظور ارزیابی دقت طبقه‌بندی مبتنی بر پارامترهای آماری حاصل از روش پیشنهادی، دقت کلی و ضریب کاپا با استفاده از نقاط چک زمینی برآورد گردید. همچنین تعدادی نمونه آموزشی برای کلاس‌های مورد نظر به کمک عامل خبره و بر اساس دانش فنی و بهره‌گیری از اطلاعات جانبی مانند تصاویر Google Earth، اخذ گردید. از این داده‌ها نیز برای طبقه‌بندی تصویر مورد نظر استفاده شد و نتایج آن با نتایج طبقه‌بندی حاصل از روش پیشنهادی مقایسه گردید.



شکل ۲: روش پیشنهادی برای استخراج پارامترهای آماری مورد استفاده در طبقه‌بندی

۳- پیاده‌سازی

الگوریتم پیشنهادی برای استخراج پارامترهای آماری بر روی تصویر +ETM و با نمونه‌های تصادفی ۱۰۰۰ عضوی پیاده‌سازی گردید. این الگوریتم پس از ۲۳۴ مرحله تکرار (اخذ ۲۳۴ نمونه تصادفی ۱۰۰۰ عضوی) به پایداری رسیده و پارامترهای آماری خوشه‌ها برآورد گردید. بر اساس پارامترهای استخراج شده، طبقه‌بندی‌های بیشترین شباهت[۳] (MLC) و ماشین‌های بردار پشتیبان[۴] (SVM) بر روی تصویر اعمال شد. همچنین از نمونه‌های آموزشی اخذ شده توسط عامل خبره نیز برای طبقه‌بندی استفاده گردید (شکل ۳).



الف: بخشی از تصویر منطقه ب: طبقه‌بندی با پارامترهای بهینه پیشنهادی ج: طبقه‌بندی با نمونه آموزشی عامل خبره

شکل ۳: مقایسه نتایج حاصل از طبقه‌بندی‌کننده MLC

با استفاده از نقاط چک زمینی، ماتریس ابهام[۵] برای طبقه‌بندی حاصل از پارامترهای آماری برآورد شده (جدول ۱) و همچنین طبقه‌بندی منتج از نمونه‌های آموزشی عامل خبره (جدول ۲) محاسبه گردید.

جدول ۱: ماتریس ابهام طبقه‌بندی MLC و SVM حاصل از پارامترهای آماری برآورد شده از روش پیشنهادی

کلاس	پیکسل‌های مرجع			
	آب	پوشش گیاهی	زمین بایر	مجم ع
آب	۷۸	۱۳	۰	۹۱
پوشش گیاهی	۰	۱۳۰	۸	۱۴۵
زمین بایر	۷	۱	۱۲۰	۱۲۱
مجموع	۸۵	۱۴۴	۱۲۸	۳۵۷
دقت کلی MLC= ۹۱/۸۸		ضریب کاپا= ۰/۸۷		

کلاس	پیکسل‌های مرجع			
	آب	پوشش گیاهی	زمین بایر	مجموع
آب	۷۸	۱۰	۰	۸۸
پوشش گیاهی	۷	۱۳۳	۷	۱۴۷
زمین بایر	۰	۱	۱۲۱	۱۲۲
مجموع	۸۵	۱۴۴	۱۲۸	۳۵۷
دقت کلی SVM= ۹۲/۹۹		ضریب کاپا= ۰/۸۹		

جدول ۲: ماتریس ابهام طبقه‌بندی MLC و SVM حاصل از نمونه‌های آموزشی عامل خبره

کلاس	پیکسل‌های مرجع			
	آب	پوشش گیاهی	زمین بایر	مجموع
آب	۶۷	۰	۰	۶۷
پوشش گیاهی	۰	۱۳۳	۷	۱۴۰
زمین بایر	۱۸	۱۱	۱۲۱	۱۵۰
مجموع	۸۵	۱۴۴	۱۲۸	۳۵۷
دقت کلی MLC= ۸۹/۹۲		ضریب کاپا= ۰/۸۴		

کلاس	پیکسل‌های مرجع			
	آب	پوشش گیاهی	زمین بایر	مجموع
آب	۷۸	۰	۰	۷۸
پوشش گیاهی	۷	۱۴۳	۱۱	۱۶۱
زمین بایر	۰	۱	۱۱۷	۱۱۸
مجموع	۸۵	۱۴۴	۱۲۸	۳۵۷
دقت کلی SVM= ۹۴/۶۷		ضریب کاپا= ۰/۹۱		

جهت بررسی میزان تطابق نقشه‌های کاربری به دست آمده از روش پیشنهادی و نمونه‌های آموزشی اخذ شده توسط عامل خبره، ماتریس تغییرات برای هر دو طبقه‌بندی MLC و SVM محاسبه گردید (جدول ۳). بر این اساس، نقشه‌های کاربری حاصل از دو روش و بر اساس طبقه‌بندی کننده‌های MLC و SVM به ترتیب ۸۸/۶۲ و ۸۷/۴۹ درصد باهم تطابق دارند.

جدول ۳: ماتریس تغییرات نقشه‌های کاربری به دست آمده از روش پیشنهادی و نمونه‌های آموزشی اخذ شده توسط عامل خبره

MLC (پارامترهای آماری برآورد شده)	MLC (داده‌های آموزشی عامل خبره)			
		آب	پوشش گیاهی	زمین بایر
	آب	۴۴۷۰	۷۱۸	۴۴۰
	پوشش گیاهی	۰	۵۶۶۳۵	۱۰۴۸۷
	زمین بایر	۰	۲۶۳۳	۵۰۱۱۳
کل کلاس		۵۶۲۸	۶۷۱۲۲	۵۲۷۴۶
تغییرات کلاس		۱۱۵۸	۱۰۴۸۷	۲۶۳۳

SVM (پارامترهای آماری برآورد شده)	SVM (داده‌های آموزشی عامل خبره)			
		آب	پوشش گیاهی	زمین بایر
	آب	۴۷۵۶	۸۸۰	۰
	پوشش گیاهی	۰	۶۳۵۷۳	۸
	زمین بایر	۰	۱۴۸۱۳	۴۱۴۶۶
کل کلاس		۴۷۵۶	۷۹۲۶۶	۴۱۴۷۴
تغییرات کلاس		۰	۱۵۶۹۳	۸

۴- بحث و نتیجه‌گیری

طبقه‌بندی کننده‌های نظارت‌شده به منظور برآورد پارامترهای آماری کلاس‌ها از نمونه‌های آموزشی بهره می‌گیرند. علاوه بر پرهزینه و زمان‌بر بودن فرآیند اخذ نمونه‌های آموزشی بر اساس مشاهدات زمینی، در برخی موارد دسترسی به منطقه مورد نظر امکان‌پذیر نبوده و همچنین در مطالعاتی مانند تغییرات کاربری اراضی، در اغلب موارد نمونه‌های آموزشی تصاویر چندزمانه در دسترس نمی‌باشد. از طرفی، تفسیر بصری توسط عامل خبره نیز محدودیت‌ها و عدم‌قطعیت‌های خاص خود را دارد. در تحقیق حاضر، از الگوریتم خوشه‌بندی k-means برای برآورد پارامترهای آماری کلاس‌ها استفاده گردید. بدین منظور، نخست عملکرد الگوریتم k-means در محاسبه پارمترهای آماری خوشه‌ها بررسی گردید. بر این اساس، نشان داده شد که این الگوریتم در تکرارهای مختلف جواب منحصر به فردی نداشته و از این روی،

مقادیر پارامترهای حاصل از این الگوریتم با نوسانات شدیدی همراه است. جهت رفع این مشکل، میانگین پارامترهای به دست آمده از k-means به ازای نمونه‌های ۱۰۰۰ عضوی از تصویر به عنوان پارامترهای بهینه کلاس‌ها انتخاب گردید. تعداد نمونه‌های تصادفی تا زمانی که پارامترهای بهینه به پایداری برسند افزایش یافته و الگوریتم k-means بر روی آنها اعمال گردید. در تصویر +ETM مورد استفاده در این تحقیق با انتخاب ۲۳۴ نمونه تصادفی ۱۰۰۰ عضوی از تصویر، پارامترهای آماری خوشه‌ها برآورد گردید. در نهایت، پارامترهای برآورد شده به عنوان جایگزینی برای نمونه‌های آموزشی در طبقه‌بندی نظارت‌شده مورد استفاده قرار گرفت. بر اساس نتایج، طبقه‌بندی MLC و SVM دقت ۹۱/۸۸ و ۹۲/۹۹ را با استفاده از پارامترهای آماری برآورد شده، به همراه داشته است. همچنین در مقایسه با نتایج طبقه‌بندی به دست آمده از نمونه‌های آموزشی اخذ شده توسط عامل خبره، تفاوت بسیار ناچیز بوده و در طبقه‌بندی MLC دقت طبقه‌بندی مبتنی بر پارامترهای برآورد شده اندکی هم بیشتر می‌باشد.

با توجه به این که در روش پیشنهادی تعداد زیادی نمونه تصادفی از تصویر انتخاب شده و میانگین پارامترهای خوشه‌ها در تکرارهای مختلف به عنوان پارامترهای بهینه در نظر گرفته می‌شود، مرکز (میانگین) و انحراف معیار خوشه‌ها در باندهای مختلف پایدار شده و استفاده از آنها در طبقه‌بندی نظارت‌شده دقت مناسبی را به همراه دارد. از این‌رو، صرف نظر از یافتن پارامترهای آماری مورد نیاز برای طبقه‌بندی کننده‌های نظارت‌شده، این روش می‌تواند به عنوان یک الگوریتم خوشه‌بندی پایدار مورد توجه قرار گیرد. از طرف دیگر، در این روش تنها لازم است که تعداد کلاس‌ها به الگوریتم معرفی شود و در ادامه در یک فرآیند خودکار، پارامترهای آماری کلاس‌ها استخراج می‌گردد. روش پیشنهادی می‌تواند مشکل اخذ داده‌های آموزشی به روش بصری را مرتفع نماید. البته لازم به یادآوری است که دانش فنی و شناخت نسبی از منطقه برای تشخیص صحیح تعداد کلاس‌های موجود در صحنه تصویر هم‌چنان مورد نیاز است و از این روی، نمی‌توان این روش را یک روش تمام خودکار دانست. به عنوان یک افق تحقیقاتی در این زمینه، بررسی امکان استفاده از الگوریتم‌هایی نظیر ISODATA را برای تشخیص خودکار تعداد کلاس‌ها می‌تواند پیشنهاد گردد.

۵- پانوشت

1. Feature Space
2. Statistic Parameters
3. Maximum Likelihood Classifier
4. Support Vector Machines
5. Confusion Matrix

منابع و مراجع

[1] Panchal, V.K., Singh, P., Kaur, N., Kundra, H., 2009. *'Biogeography based Satellite Image Classification'*, (IJCSIS) International Journal of Computer Science and Information Security, Vol. 6 , No.2.

[2] Long, W., Srihann, S., 2004. *'Land Cover Classification of SSC Image: Unsupervised and Supervised Classification Using ERDAS Imagine'*, in IEEE International Geoscience and Remote Sensing Symposium (2004) Volume 4, IEEE Computer Society Press, pp.2707- 2712.

[3] Omkar S.N., Manoj, K., Dipti, M., 2007. *'Urban Satellite Image Classification using Biologically Inspired Techniques'*, in IEEE International Symposium on Industrial Electronics(2007), IEEE Computer Society Press, pp. 1767-1772.

[4] Liu, X.H., Skidmore, A.K., and Oosten, V.H., 2002. *'Integration of Classification Methods for Improvement of Land-Cover Map Accuracy'*, ISPRS Journal of Photogrammetry& Remote Sensing, 56, pp. 257-268.

[5] Richards, J.A., Jia, X., 2006. *'Remote Sensing Digital Image Analysis'*, 4th Edition. Springer-Verlag Berlin Heidelberg.

[6] Theodoridis, S., Koutroumbas, K., 2003. *'Pattern Recognition'*, 2nd Edition. ACADEMIC PRESS, An imprint of Elsevier.

[7] Jensen, J.R, 2007. *'Remote Sensing of the Environment: an Earth Resource Perspective'*, 2nd ed. Upper Saddle River, NJ, Prentice Hall. pp. 592.

[8] Mather, P. M., 2004. *Computer Processing of Remotely-Sensed Images: An Introduction*. Third Edition. London: John Wiley & Sons Ltd.

Semi-Automatic Extraction of the Statistical Parameters of the Interest Classes based on the Image Clustering

Mojtaba Jannati [a, *], Milad Niroumand Jadidi, Mohammad Javad Valadan Zoej [a], Ali Mohammadzadeh [a], Alireza Safdarinezhad [a]

[a] Dept. of Remote Sensing Engineering, Geomatics Engineering Faculty, K. N. Toosi University of Technology, ValiAsr Street, Mirdamad Cross, P.C. 19967-15433, Tehran, Iran - (m.jannati, niroumand, safdari_nezhad)@sina.kntu.ac.ir, (valadanzouj, a_mohammadzadeh)@kntu.ac.ir

Abstract

Acquisition of the training data could be considered as the most important stage of the supervised classification process. Direct field observations and also the available reference data are known as reliable sources of gathering this data. This process is usually time/cost consuming and also in some cases, it is not practical due to some limitations (e.g. unreachability of the study area, using the archived data). So, the visual interpretation of an expert user is often considered as a solution to acquire at least some part of the training data. However, this solution is followed by its own uncertainties and also is time consuming. In this research, a semi-automated approach is proposed to estimation of the statistical parameters of the classes in order to use in the supervised classification. In this way, the instability problem of the statistical parameters obtained from the k-means clustering algorithm is solved in an iterative process and then the optimal parameters for each class are estimated. So, the resultant parameters are used instead of the training data. According to the results, the overall accuracies of the maximum likelihood and Support Vector Machines (SVM) algorithms based on the estimated parameters are determined 91.88% and 92.99% and the Kappa coefficient equals to 0.87 and 0.89 respectively. On the other hand, some training data is acquired by visual interpretation of an expert user and the accuracy of classification which done based on this data is compared with results obtained based on the estimated parameters. In conclusion, the estimated parameters lead to an accurate supervised classification. The proposed method could be considered as a semi-automated procedure for obtaining the statistical parameters which just the introduction of the number of classes is required.

Keywords: Supervised Classification, Training Data, Statistical Parameters, Clustering