



نسخه از دور

GIS ایران



سنجش از دور و GIS ایران سال یازدهم، شماره سوم، پاییز ۱۳۹۸
Iranian Remote Sensing & GIS Vol.11, No. 3, Autumn 2019

۴۲-۲۱

تشخیص سازه‌های ساخت بشر در تصاویر هوایی با استفاده از ویژگی‌های آماری مبتنی بر رنگ و یادگیری ماشین

ناصر فرج‌زاده^۱ و مهدی هاشم‌زاده^۲

۱-دانشیار دانشکده فناوری اطلاعات و مهندسی کامپیوتر، دانشگاه شهید مدنی آذربایجان، ایران

۲-دانشیار دانشکده فناوری اطلاعات و مهندسی کامپیوتر، دانشگاه شهید مدنی آذربایجان، ایران

تاریخ پذیرش مقاله: ۱۳۹۸/۰۲/۱۷

تاریخ دریافت مقاله: ۱۳۹۷/۰۸/۱۳



چکیده

تصاویر هوایی ثبت‌شده توسط ماهواره‌ها و یا پهپادها، معمولاً شامل نواحی مربوط به منابع طبیعی و نواحی حاوی سازه‌های ساخت بشر است. با تفکیک این نواحی از یکدیگر، قادر خواهیم بود اطلاعات مهمی نظیر آرایش ساختاری سطوح و شکل آنها را استخراج و نقشه‌های هوایی برچسب‌دار ایجاد کنیم. دستیابی به چنین اطلاعاتی می‌تواند کاربردهای بسیار مفیدی در زمینه‌های نظامی، شهری و زیست محیطی داشته باشد. از آنجایی که پردازش حجم عظیمی از تصاویر به‌دست آمده از ماهواره‌ها و پهپادها بصورت دستی امکان‌پذیر نیست، لذا استفاده از روش‌های خودکار مبتنی بر هوش مصنوعی در این حوزه، بسیار مورد توجه قرار گرفته است. تاکنون پژوهش‌های متعددی در این‌باره انجام شده است که از اهم آنها می‌توان به شناسایی ساختمان‌ها، وسایل نقلیه، جاده‌ها و همچنین تشخیص ساختار پوشش گیاهی در تصاویر هوایی اشاره کرد. در این مقاله، قصد داریم با معرفی مجموعه‌ای از ویژگی‌های آماری مبتنی بر رنگ که به سادگی قابل استخراج از تصاویر هوایی هستند و با استفاده از یک مدل مبتنی بر یادگیری، راهکاری کارآمد برای تشخیص دقیق و سریع سازه‌های ساخت بشر و تفکیک آنها از منابع طبیعی ارائه دهیم. نتایج آزمایش‌های انجام گرفته بر روی بانک تصاویر ماساچوست که به‌صورت عمومی، قابل دسترس است، نشان دهنده دقت مناسب و سرعت عمل بالای راهکار پیشنهادی است. به‌طوری‌که، دقت و میانگین سرعت پردازش به‌دست آمده به ترتیب برابر با ۹۰/۰۷٪ و ۰/۹۶ ثانیه است.

کلیدواژه‌ها: تصویر هوایی، منابع طبیعی، سازه‌های ساخت بشر، یادگیری ماشین، ویژگی‌های آماری

*نویسنده عهده‌دار مکاتبات: تبریز، ۳۵ کیلومتری جاده تبریز مراغه، دانشگاه شهید مدنی آذربایجان، دانشکده فناوری اطلاعات و مهندسی کامپیوتر، کدپستی: ۵۳۷۵۱۷۱۳۷۹، تلفن

همراه: ۰۹۳۰۸۱۴۹۶۰۰

(۱) مقدمه

ساختمان‌ها، اتصال آنها به هم و وجود موانعی در کنار آنها نظیر درخت، مساله شناسایی با مشکلات جدی روبرو می‌شود (Burochin, Vallet et al. 2014). ذکر این نکته ضروری است که براساس اقلیم‌ها و نوع ساختار شهری موجود در هر منطقه، مطلوب آن است که بتوان برای تمامی ساختارهای شهری این عمل را انجام داد. در اکثر شهرها، ساختمان‌ها به هم چسبیده‌اند که در چنین شرایطی، معمولاً عمل تشخیص مشکل‌تر می‌شود. به‌عنوان مثال، شناسایی و استخراج جاده از تصاویر هوایی در کاربردهایی نظیر نرم‌افزارهای کمکراننده و وسایل نقلیه هوشمند بسیار مهم است. همچنین شناسایی و شمارش تعداد خودروها در تصاویر هوایی می‌تواند نقش مهمی در مدیریت ترافیک شهری و بین شهری ایفا نماید.

در کاربردهای متنوع، مبتنی بر برچسب‌گذاری و تفسیر تصاویر هوایی، استخراج ساختمان‌ها، جاده‌ها و سایر اشیا معمولاً توسط انسان و به‌صورت دستی انجام می‌شود. بدیهی است که چنین فرآیندی می‌تواند بسیار پرهزینه و وقت‌گیر باشد. به‌منظور خودکارسازی این عمل، تلاش‌های زیادی صورت گرفته است. در بسیاری از روش‌های متداول، از ویژگی‌های محلی برای طبقه‌بندی پیکسل‌ها و یا بخش‌ها استفاده شده است (Cheng, Han et al. 2016, Hussnain, Elberink et al. 2016, Nagarajan, Schenk et al. 2016, Xia, Chen et al. 2016). با این حال، در اکثر این روش‌ها، داشتن دانش قبلی درباره ظاهر شیء ضروری بوده و در کاربردهای بلادرنگ معمولاً حصول چنین دانشی غیر ممکن است (Cheng, Han et al. 2016). همچنین، در برخی مواقع حتی با وجود دانش قبلی، دقت بازشناسایی این روش‌ها در حضور چالش‌های موجود در تصاویر هوایی، به‌طور قابل ملاحظه‌ای کاهش پیدا می‌کند. از این‌رو، طراحی و ارایه راهکاری که بتواند حجم انبوهی از تصاویر هوایی را در مدت زمان کم و با دقت بالا بدون نیاز به هیچ‌گونه دانش پیشین برچسب‌گذاری کند، یکی از نیازهای اساسی در کاربردهای مبتنی بر تفسیر تصاویر هوایی به‌شمار می‌رود.

امروزه با افزایش دسترسی به پهپادها و سهولت گردآوری تصاویر هوایی ثبت شده از سامانه‌هایی نظیر Google Earth، نقش چنین تصاویری در جمع‌آوری و استخراج اطلاعات سودمند پرننگ و در برخی موارد، حیاتی شده است. در اکثر این روش‌ها، تشخیص و استخراج محل قرارگیری اشیا در تصاویر هوایی که معمولاً شامل نواحی مربوط به منابع طبیعی و نواحی حاوی سازه‌های ساخت بشر است، مهم‌ترین بخش از کار به‌حساب می‌آید. به‌طوری‌که، هر قدر کارایی این بخش بیشتر باشد، فهم بهتر و موثرتری از تصاویر برای پردازش‌های بعدی حاصل می‌شود (Cao and Yang 2007).

در تحلیل و درک چنین تصاویری، شناسایی اشیا ساخت بشر نظیر ساختمان‌ها (سجادی و پارسیان، ۱۳۹۶)، خودروها و شمارش تعداد آنها (Moranduzzo, Melgani et al. 2014)، جاده‌ها (Tan, Wang et al. 2008)، شناسایی نواحی طبیعی و تشخیص پوشش گیاهی آنها (Wei, Xin et al. 2010)، استخراج اطلاعات مهم در مورد آرایش ساختاری از سطوح و ایجاد نقشه‌های هوایی (Thuy 2009)، از مسائل بسیار مهم در حوزه سنجش از دور هستند. به‌عنوان مثال، ایجاد مدل‌های شهری سه‌بعدی با استفاده از تصاویر هوایی، یکی از موضوعات بسیار پرکاربرد و در عین حال چالش برانگیز در این حوزه به‌شمار می‌رود. زیرا علاوه بر اطلاعات ارتفاعی که در این مدل‌سازی‌ها لازم است، اطلاعات سطحی نظیر سقف ساختمان‌ها و معابر نیز می‌بایست از تصاویر هوایی استخراج شده و به کمک اطلاعات جانبی نظیر تصاویر زمینی، مدل سه‌بعدی مورد نظر تشکیل شود (Fischer, Kolbe et al. 1998). برای ایجاد ساختمان‌ها در این مدل شهری، ابتدا باید آنها را شناسایی نمود. سپس به بازسازی و درنهایت به واقعی‌سازی آنها پرداخت. بیشتر پژوهش‌های انجام صورت گرفته برای تشخیص کلی نواحی ساختمانی با فرض ساده و جدا از هم بودن آن نواحی بوده و با پیچیده‌تر شدن شکل

کار آن را استفاده از ویژگی‌های آماری مبتنی بر رنگ و یادگیری ماشین تشکیل داده است، نشان از دقت و سرعت عمل بالای آن دارد.

در ادامه، در بخش ۲ پیشینه پژوهش و روش‌های موجود در حوزه تشخیص و شناسایی اشیای ساخت بشر در تصاویر هوایی ارائه می‌شوند، در بخش ۳ جزئیات راهکار پیشنهادی توضیح داده می‌شود، در بخش ۴ نتایج آزمایش‌های انجام گرفته برای ارزیابی کارایی راهکار پیشنهادی مورد بحث و بررسی قرار می‌گیرد و نهایتاً در بخش ۵ یافته‌های کلی پژوهش و مسیرهای ممکن برای توسعه و مطالعه بیشتر ارائه می‌شوند.

۲) پیشینه پژوهش

مطالعات و کارهای کاربردی انجام‌شده با موضوع تشخیص سازه‌های ساخت بشر در تصاویر هوایی را می‌توان در سه گروه (۱) تشخیص ساختمان‌ها، (۲) تشخیص خودرو و وسایل نقلیه و (۳) تشخیص جاده و مسیرهای ارتباطی، دسته‌بندی کرد. در اغلب روش‌های مطرح شده، استخراج و ارزیابی ویژگی‌های بافت تصویر مبنای تشخیص و طبقه‌بندی است و از این‌رو، پژوهش‌های مهم و متعددی در این‌باره صورت گرفته است. با توجه به این‌که هدف اصلی این مقاله، ارائه یک روش استخراج ویژگی برای توصیف بهتر بافت در تصاویر هوایی است، لذا مرور روش‌های پیشین با این رویکرد و به ترتیب زمان ارائه پژوهش‌های انجام می‌گیرد.

گروئن که یکی از محققین پیشگام در حوزه شناسایی خودکار سازه‌های ساخت بشر در تصاویر هوایی به شمار می‌رود، یک روش نیمه خودکار برای استخراج جاده ارائه کرد که ترکیبی از عمل تجزیه موجک^۱ و یک الگوریتم استخراج ویژگی‌های خطی بر اساس برنامه‌نویسی پویا بود (Gruen, Li et al. 1995). در این روش بعد از این‌که چند نقطه بذر توسط کاربر انتخاب می‌شود، عمل

با توجه به مقدمه فوق، می‌توان مساله تشخیص سازه‌های ساخت بشر در تصاویر هوایی را به دو دلیل عمده چالش برانگیز دانست: (۱) پیچیدگی ظاهری-بافتی و گوناگونی سازه‌ها و اشیای در تصاویر هوایی و (۲) بار محاسباتی بسیار بالا برای پردازش حجم انبوهی از تصاویر (Cheng and Han 2016). از آنجایی‌که یک معیار کلی برای استخراج و بازشناسایی تمامی اشیای موجود در تصاویر وجود ندارد، لذا روش‌های پیشنهاد شده برای حل چالش اول معمولاً با هدف شناسایی دسته خاصی از اشیای نظیر ساختمان، خودرو و یا جاده مطرح شده‌اند (Cheng, Han et al. 2016). دسته دوم با تاکید بر کاهش زمان پردازش دنبال راهکارهایی هستند که با رعایت حداقل دقت تفکیک قابل قبول، زمان پردازش را کاهش و سرعت تفکیک را بالا ببرند. به‌عبارتی، در روش‌های قرار گرفته در این دسته، با استفاده از ویژگی‌هایی که به سادگی و بدون نیاز به پردازش‌های پیچیده قابل استخراج هستند، سعی در توصیف اشیای و یا بافت نواحی مورد نظر در تصویر می‌شود (Cai, Chen et al. 2011). همچنین، سعی می‌شود تا با کاهش طول بردار توصیف‌کننده توسط روش‌هایی نظیر تحلیل مولفه‌های اصلی^۱ (Cunningham and Ghahramani 2015)، متوسط سرعت پردازش (تفکیک) را افزایش دهند (Cheng, Han et al. 2016).

در این مقاله، با هدف افزایش سرعت عمل تفکیک و رسیدن به دقت قابل قبول، مجموعه‌ای جدید از ویژگی‌های آماری برای توصیف تصاویر هوایی و طبقه‌بندی نواحی مختلف تصویر به مناطق طبیعی و غیرطبیعی (ساخت بشر) معرفی می‌شوند. انگیزه ارائه این راهکار بر این استوار است که ترکیب چند مولفه آماری مبتنی بر رنگ که به‌راحتی و با سرعت زیادی قابل استخراج از تصاویر هستند و استفاده از یادگیری ماشین برای ساخت مدل طبقه‌بندی کننده، می‌تواند باعث بهبود عملکرد تشخیص سازه‌های ساخت بشر در تصاویر هوایی شود.

ارزیابی راهکار پیشنهادی در این مقاله، که اساس

1. Principle Component Analysis (PCA)

2. Wavelet

لاو^۵، ماتریس مجاورت^۶، طول اولیه^۷، فرکانس لبه^۸ و همبستگی^۹ را باهم مقایسه کردند. با توجه به نتایج بدست آمده، روش لاو و روش ماتریس مجاورت از کارایی بهتری نسبت به بقیه روش‌ها برخوردار بودند. رویز و همکارش (Ruiz, Fdez-Sarría et al. 2004) برای تجزیه و تحلیل روش‌های توصیف‌کننده بافت، از تصاویر هوایی با وضوح بالا که از محیط‌های شهری و جنگلی تهیه شده بودند، استفاده کردند. در پژوهش آنها چهار روش برای تجزیه و تحلیل بافت و استخراج ویژگی‌ها بررسی شده است: (۱) روش آماری بر اساس سطح خاکستری ماتریس مجاورت، (۲) فیلتر انرژی و عامل لبه^۱، (۳) فیلتر گابور و (۴) موجک. آنها با ارزیابی این روش‌ها به این نتیجه رسیدند که نمی‌توان به یک روش عمومی کارآمد استخراج ویژگی، برای توصیف تصاویر هوایی دست یافت، و ترکیب چند روش می‌تواند موجب بهبود کارایی شود.

هینز (Hinz 2003) یک روش جدید برای تشخیص خودرو و تعداد آن به صورت خودکار در مقیاس بزرگی از تصاویر هوایی، معرفی کرد. در این روش، رنگ خودرو، شدت روشنایی شیشه جلو و شدت سایه اطراف خودرو به عنوان ویژگی‌های رادیومتری جهت توصیف بافت در نظر گرفته شدند. در پژوهشی که توسط فاگویر و همکارانش (Fauqueur, Kingsbury et al. 2005) صورت گرفت، توصیف بافت در تصاویر هوایی براساس ویژگی‌های رنگی، بافتی و ساختاری انجام شد. هدف آنها ارزیابی روشی بود تا بتواند به‌طور معنایی و بر اساس ظاهر بصرفه بافت، عملیات برچسپ‌گذاری را بصورت خودکار انجام دهد. باگواتی و همکارانش (Bhagavathy, Manjunath et al. 2006) از نقوش بافت و ویژگی‌های فضایی برای مدل‌سازی و تشخیص بندر، زمین گلف، خانه مسکونی و پارکینگ استفاده کردند.

تشخیص جاده صورت می‌گیرد. یک‌سال بعد، برزوهار و همکارش (Barzohar, Cooper et al. 1996) یک روش هندسی جدید برای تشخیص جاده‌ها در تصاویر هوایی ارائه دادند. تاکید این پژوهش بر روی استخراج خودکار جاده‌های اصلی در موقعیت‌های انحنا جاده بر اساس عرض، شدت نور و لبه‌های جاده در تصویر بود. در روش پیشنهادی آنها از توزیع گیبس^۱ استفاده شده است.

ما و همکارش (Ma and Manjunath 1998)، از توصیف‌گر موجک گابور^۲ و نقشه‌های خود سازمان‌ده^۳ برای طبقه‌بندی عکس‌های هوایی استفاده کردند. در پژوهشی همزمان، لین و همکاران (Lin, Nevatia et al. 1998) روشی برای شناسایی ساختمان‌ها و اشیای سه‌بعدی در تصاویر هوایی با استفاده از شدت روشنایی سطح ارائه شد. هدف اصلی در روش پیشنهادی آنها، تشخیص ساختمان‌های نظامی و صنعتی با فرض مسطح‌بودن سقف این نوع ساختمان‌ها بود. با این فرض و با استفاده از تصاویر سه‌بعدی، عمل قطعه‌بندی تصاویر و استخراج ویژگی‌ها با استفاده از روش خطی^۴ انجام گرفت. در پژوهشی دیگر فیشر یک روش مبتنی بر مدل‌های دوبعدی و سه‌بعدی ارائه کرد که به‌صورت خودکار ساختمان‌های سه‌بعدی را از تصاویر هوایی استخراج می‌کرد (Fischer, Kolbe et al. 1998). ویژگی‌های استخراج شده از ساختمان شامل ویژگی‌های خطوط (افقی، عمودی و مورب) و ویژگی‌های نقاط مختلف از ساختمان (مثلاً از سقف و دیوارها پارامترهایی مثل راستا، عرض، طول و ارتفاع) بود.

باگواتی و همکارانش (Bhagavathy, Newsam et al. 2002) بردارهای همگن را با استفاده از فیلتر موجک گابور استخراج کرده و با بهره‌گیری از روش مخروط گاوسی ویژگی‌هایی را جهت شناسایی سازه‌های ساخت دست بشر ارائه کردند. نتایج نشان داد که این روش قادر به تجزیه و تحلیل معناداری از نقوش بافت بوده و نسبت به چرخش مقاوم است. شارما و همکارش (Sharma and Singh 2001) پنج روش توصیف بافت

1. Gibbs distributions
2. Gabor wavelet
3. Self-Organizing Maps (SOM)
4. Linear feature extraction
5. Law's Method
6. Co-occurrence matrix method
7. Primitive-length
8. Edge frequency
9. Auto-correlation
10. Edgeness

هوایی می‌شوند. با توجه به کارایی بسیار خوبی که روش‌های استخراج ویژگی خودکار مبتنی بر یادگیری عمیق در حوزه بینایی ماشین از خود نشان داده‌اند، در چند سال اخیر استفاده از این روش بصورت گسترده‌ای مورد قبول واقع شده است. لذا پنااتی و همکارانش (Penatti, Nogueira et al. 2015) کارایی روش استخراج خودکار ویژگی‌ها را به کمک یادگیری عمیق^۸ آزمایش کرده و به این نتیجه رسیده‌اند که استفاده از این روش از عملکرد بسیار مناسبی در توصیف تصاویر هوایی برخوردار است. علاوه بر این، آنها دریافتند که روش‌های استخراج ویژگی‌های سطح پایین مبتنی بر رنگ نظیر روش طبقه‌بندی پیکسل حاشیه‌ای و درونی^۹ (Stehling, Nascimento et al. 2002) کارایی بهتری در سنجش از راه دور تصاویر، از خود نشان می‌دهند.

در دسته‌ای دیگر از مقالات اخیر (Mnih 2013, Saito and Aoki 2015, Saito, Yamashita et al. 2016)، روش‌های مبتنی بر یادگیری عمیق در تشخیص ساختمان‌ها و جاده‌ها به کار برده شده است که نتایج گزارش شده حاکی از کارایی بسیار بالای این روش‌ها از نظر دقت است. قابل ذکر است که مزیت اصلی یادگیری عمیق نسبت به سایر روش‌های یادگیری در استخراج خودکار ویژگی‌ها در آن است. مرحله استخراج ویژگی‌ها که مهمترین و تاثیر گذارترین بخش هر سیستم خودکار محسوب می‌شود، نیاز به دقت و وسواس ویژه‌ای دارد تا با انتخاب روش استخراج ویژگی مناسب، به کارایی دلخواه دست پیدا کرد. این مرحله، در روش‌های مبتنی بر یادگیری عمیق با استفاده از شبکه‌های عصبی چندلایه و بدون دخالت و یا مهندسی کردن روش استخراج ویژگی، انجام می‌پذیرد. با توجه به ساختار

اوزدمیر و همکارش آکسای (Ozdemir and Aksoy 2010) از ترکیب مدل کیسه کلمات^۱ و استخراج گراف تکرارشونده^۲ برای ساخت ویژگی‌های سطح بالاتر در راستای توصیف بافت در تصاویر هوایی استفاده کردند. وانگ ویی (Wei, Xin et al. 2010) یک روش تشخیص اشیای ساخت بشر را در تصاویر هوایی بر اساس الگوریتم قطعه‌بندی آب‌پخش‌ن تغییر یافته^۳ ارائه داد. در این روش، استفاده از مبدل کانتورلت^۴ بدون استفاده از نمونه‌برداری^۵ برای توصیف بافت، پیشنهاد شده است. لی و همکارش (Li and Itti 2011) به بررسی روش خودکار شناسایی و طبقه‌بندی اهداف در تصاویر هوایی با وضوح بالا که متکی به ویژگی‌های آماری و سطح پایین بیولوژیک بود، پرداختند. بورشین و همکارانش (Burochin, Vallet et al. 2014) روشی برای شناسایی ساختمان‌هایی با نمای کور در تصاویر هوایی را با استفاده از دوربین‌های قوی در مقیاس دسی‌متر ارائه کرده است. ساختمان‌های کور، ساختمان‌هایی هستند که هیچ دهانه‌ای مثل در یا پنجره برای تشخیص آنها وجود ندارد. در روش ارائه شده، استفاده از ویژگی‌های محلی و ویژگی‌های سراسری مطرح شده است: در ویژگی‌های محلی بر رادیومتری عمومی و اطلاعات هندسی درباره بافت شهری و ساختمان‌ها تکیه شده است و در ویژگی‌های سراسری شدت نور محلی تصاویر، مورد استفاده قرار گرفته است. در این روش پس از توصیف خارجی ساختمان با استفاده از ویژگی‌های محلی و سراسری، از ماشین بردار پشتیبان^۶ برای طبقه‌بندی و بازشناسایی بهره گرفته شده است.

در مطالعه‌ای که اخیراً صورت گرفته است، چیردیات و همکارش (Cheryadat and Sensing 2014) روشی خودکار برای استخراج ویژگی‌های بافت، معرفی کرده‌اند. در روش آنها، ویژگی‌های سطح پایین نظیر شدت روشنایی پیکسل‌ها استخراج شده و سپس این ویژگی‌های سطح پایین توسط توابع پایه‌ای که بصورت خودکار محاسبه می‌شوند، تبدیل به ماتریس‌های خلوت^۷ جهت بازنمایی بافت در تصاویر

1. Bag-of-words
2. Frequent subgraph mining
3. Modified watershed segmentation
4. Contourlet
5. Non-subsampled contourlet transform
6. Support Vector Machines (SVM)
7. Spars representation
8. Deep learning
9. Border/Interior pixel classification

ویژگی‌های متعددی از قطعات استخراج می‌شود. ویژگی‌های استخراج‌شده، به یک بخش یادگیری ماشین وارد شده تا یک مدل طبقه‌بندی‌کننده تصاویر ساخته شود. مدل آموزش‌دیده شده می‌تواند برای طبقه‌بندی تصاویر دیگری که در آینده به سیستم وارد می‌شوند، مورد استفاده قرار گیرد. در شکل ۱، دیاگرام راهکار پیشنهادی نشان داده شده است. در ادامه، هر یک از مراحل انجام کار با جزئیات بیشتری توضیح داده می‌شود.

۳-۱) تبدیل فضای رنگ

رنگ، همواره به‌عنوان یک ویژگی بصری بسیار قوی در شناسایی اجسام به‌شمار می‌رود (Li, Chen et al. 2002). رنگ‌ها که به‌صورت ترکیبی از امواج با فرکانس‌های متفاوت در قالب سه رنگ قرمز، سبز و آبی برای انسان قابل درک هستند، فضای رنگی RGB را تشکیل می‌دهند. همچنین فضاهای رنگی دیگر نظیر YUV, YIQ, LAB, CIE LUV, CIE و HSV را می‌توان با جداسازی شدت روشنایی از اطلاعات رنگی بدست آورد. با این حال هیچ‌کدام از این فضاهای رنگی را نمی‌توان به‌عنوان یک فضای رنگی غالب در همه کاربردها به‌شمار آورد. به‌عنوان مثال، در بیشتر دستگاه‌ها از فضای رنگی RGB و در تلوزیون‌های رنگی معمولاً از فضای رنگی YIQ و یا YUV استفاده می‌شود. فضاهای رنگی CIE LAB و CIE LUV نیز به‌عنوان فضاهای رنگی همسان شناخته می‌شوند. در میان فضاهای رنگی، فضای رنگی HSV و مشتقات آن به‌طور ویژه‌ای نزدیک به نحوه درک انسان از رنگ‌ها است، از این‌رو در بیشتر سیستم‌های خودکاری که الهام‌گرفته از سیستم بینایی انسان باشند، از این فضای رنگی استفاده می‌شود (Douglas and Kirkpatrick 1999).

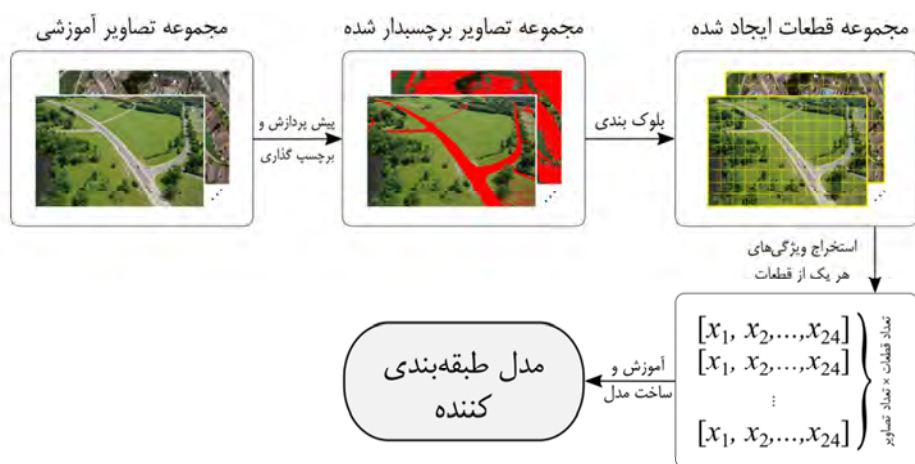
پیچیده یادگیری عمیق، فرآیند ساخت مدلی مبتنی بر آن، فوق‌العاده زمان‌بر بوده و همچنین نیاز به فراهم بودن داده‌های آموزشی بسیار زیادی دارد.

مطالعه‌ی روش‌های آرایه شده در بازشناسایی سازه‌های ساخت بشر در تصاویر هوایی نشان می‌دهد که استفاده از روش‌های استخراج ویژگی و توصیف بافت در تصاویر هوایی به‌صورت مهندسی شده از سرعت بسیار بیشتری در مقایسه با روش‌های مبتنی بر یادگیری عمیق برخوردار هستند (Marcu 2016). در حالی که، دقت این روش‌ها در مقایسه با روش‌های مبتنی بر یادگیری عمیق، نسبتاً پایین است. در این پژوهش، مجموعه‌ای از ویژگی‌های آماری مبتنی بر رنگ که به سادگی قابل استخراج از تصاویر هوایی هستند، معرفی می‌شوند که بهره‌گیری از آنها در یک مدل مبتنی بر یادگیری ماشین، روشی کارآمد برای تشخیص دقیق و سریع سازه‌های ساخت بشر را مهیا می‌کند.

۳) راهکار پیشنهادی

همانطور که پیش‌تر اشاره شد، شناسایی سازه‌های ساخت بشر در تصاویر هوایی و تفکیک آنها از منابع طبیعی، کاربردهای متنوعی دارد و نقش مهمی در فهم تصاویر هوایی و شاخص‌گذاری‌های محتوایی آنها ایفا می‌کند. در راهکار پیشنهادی، مجموعه جدیدی از ویژگی‌های آماری به کار گرفته می‌شوند که علاوه بر ساده بودن، باعث افزایش دقت و سرعت تشخیص سازه‌های ساخت بشر در تصاویر هوایی می‌شوند. عملکرد کلی سیستم به این صورت است که در ابتدا مجموعه‌ای از تصاویر هوایی که حاوی منابع طبیعی و سازه‌های ساخت دست بشر هستند، جهت آموزش سیستم، جمع‌آوری می‌شود. سپس این تصاویر آموزشی توسط یک ناظر انسانی برچسب‌گذاری می‌شوند: منابع طبیعی^۱ یا سازه‌های ساخت بشر^۲. در ادامه هر یک از تصاویر به قطعاتی هم‌اندازه، بلوک‌بندی شده و سپس با استفاده از روش استخراج ویژگی پیشنهادی که مبتنی بر فضاهای رنگی RGB و HSV است،

1. Natural
2. Man-made



(الف)



(ب)

شکل ۱. دیاگرام راهکار پیشنهادی. (الف) فرآیند آموزش و (ب) فرآیند تشخیص (تفکیک).

۳-۲) استخراج ویژگی‌های آماری

اطلاعات آماری که در راهکار پیشنهادی برای توصیف بافت تصاویر، استخراج می‌شوند عبارتند از: میانگین (\bar{x}) ، میانسه (\bar{x}) ، مُد (M) و واریانس (σ^2) . پس از بلوک‌بندی یک تصویر به قطعات کوچک‌تر، این شاخص‌های آماری برای هر قطعه‌ای، با استفاده از مقادیر همه پیکسل‌های موجود در آن قطعه، محاسبه می‌شوند. این ویژگی‌ها برای هر یک از کانال‌های RGB و HSV قطعات به صورت مجزا استخراج می‌شوند.

در راهکار پیشنهادی نیز به این دلیل که سیستم بر اساس نمونه‌های برچسب زده شده توسط ناظر انسانی آموزش داده می‌شود، علاوه بر اطلاعات آماری که از فضای رنگی RGB استخراج می‌شوند، از اطلاعات آماری فضای رنگی HSV نیز برای ایجاد بردار ویژگی جهت توصیف بافت تصویر استفاده می‌شود. شکل ۲ نمونه‌هایی از تبدیل فضای رنگی از RGB به HSV را در تصاویر هوایی نشان می‌دهد.



شکل ۲. نمونه‌های از تصاویر هوایی. ستون سمت راست: فضای رنگی RGB، ستون سمت چپ: فضای رنگی HSV

زمانی که داده پرت^۱ در یک مجموعه وجود نداشته باشد، میانگین را می‌توان به‌عنوان معرف مناسبی از آن مجموعه تلقی کرد. با این فرض که شدت تغییرات بافت در سازه‌های ساخت بشر زیاد نیست، میانگین شدت نور نقاط تشکیل دهنده یک بافت را می‌توان به‌عنوان معرف عمومی آن بافت دانست. اما از آنجایی که ممکن است در برخی مناطق نظیر مناطق حاشیه‌ای اشیاء، تغییرات قابل ملاحظه‌ای در بافت وجود داشته باشد، می‌توان از میانه که نسبت به داده‌های پرت حساس نیست، به‌عنوان معرف دیگری برای بافت استفاده کرد. همچنین با استفاده از مد می‌توان شدت روشنایی غالب در یک تصویر را به‌دست آورد و به دلیل این‌که بافت سازه‌های ساخت بشر در بیشتر موارد شدت نور یکسانی خواهند داشت، می‌توان از مد تصویر برای به‌دست

از این‌رو، طول بردار ویژگی هر قطعه برابر با ۲۴ مولفه $(4 \times (3 + 3))$ خواهد بود. نحوه محاسبه این ویژگی‌ها بر اساس روابط زیر است:

$$\bar{x} = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n f^c(i, j), \quad (1)$$

$$\tilde{x} = f_p^c(i, j), \quad \forall p, q \in \{1, \dots, m \times n\}, \quad (2)$$

$$p \leq q \rightarrow f_p^c(i, j) \leq f_q^c(i, j),$$

$$p = \frac{m \times n}{2} + 1 \text{ بطوری که} \quad (3)$$

$$M = \text{Max}(\text{Frequency}(f^c(i, j))), \quad (3)$$

$$\sigma^2 = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n (\bar{x} - f^c(i, j))^2. \quad (4)$$

در روابط ۱ الی ۴، m و n بترتیب نمایانگر عرض و طول تصویر، $f^c(i, j)$ نمایانگر شدت نور یک پیکسل در مختصات (i, j) ام، c نمایانگر کانال رنگی و $i, j = 1, \dots, m \times n$ است.

1. Outlier

آوردن چنین ویژگی استفاده کرد. و بالاخره این‌که، با استفاده از **واریانس** یک تصویر، می‌توان به میزان پراکندگی شدت نور در قسمت‌های مختلف یک تصویر پی برد. هر چقدر این مقدار نزدیک به صفر باشد، نشان‌گر این نکته است که بافت یکدستی وجود دارد. از آنجایی که معمولا سازه‌های ساخت بشر در تصاویر هوایی از بافت یکدستی برخوردار هستند، لذا واریانس می‌تواند ویژگی مناسبی برای توصیف چنین بافت‌هایی باشد.

۳-۳) ساخت مدل طبقه‌بندی کننده

برای تفکیک خودکار سازه‌های ساخت بشر از منابع طبیعی، بر اساس ویژگی‌های استخراج شده، لازم است که یک مدل طبقه‌بندی کننده با استفاده از تصاویر آموزشی ایجاد شود. با توجه به این‌که در مسئله پیش رو عمل طبقه‌بندی برای دو گروه انجام می‌گیرد و به دلیل این‌که ناظر انسانی نمونه‌های آموزشی را برچسب‌گذاری کرده است، الگوریتم ساخت مدل طبقه‌بندی کننده در راهکار پیشنهادی از نوع دودویی و با ناظر^۱ است. الگوریتم‌های متعددی نظیر ماشین بردار پشتیبان (Vapnik 1999)، شبکه‌های عصبی مصنوعی، درخت‌های تصمیم‌گیر و ... برای آموزش و ایجاد مدل‌های طبقه‌بندی کننده وجود دارد که در این پژوهش، از ماشین بردار پشتیبان استفاده می‌شود (رجوع شود به بخش ۳-۴). از مجموعه داده‌های آموزشی برای آموزش ماشین بردار پشتیبان، جهت مشخص کردن رابطه $F = f(\bar{x}, \bar{x}, M, \sigma^2)$ استفاده می‌شود. F یک مقدار دوتایی دارد که مشخص می‌کند یک قطعه از تصویر ساخت دست بشر است یا خیر.

ماشین‌های بردار پشتیبان، عموماً برای مسایلی که در آنها دو گروه داده، وجود دارد، از کارایی بسیار مناسبی برخوردار هستند (Burkov 2019). اگر داده‌ها به صورت خطی تفکیک‌پذیر باشند، این الگوریتم سعی خواهد کرد ماشین‌های خطی که بتوانند در قالب یک ابر صفحه، داده‌ها را بدون خطا و با حداکثر فاصله میان صفحه و نزدیکترین نقاط آموزشی (بردارهای پشتیبان)

تفکیک کنند، ایجاد نماید. زمانی که داده‌ها به صورت خطی تفکیک‌پذیر نباشند، می‌توان با استفاده از نگاشت‌های غیرخطی، داده‌ها را قبل از فرآیند ایجاد بردارهای پشتیبان، به فضایی با ابعاد بیشتر، با این امید که داده‌ها در فضای با ابعاد بیشتر، به صورت خطی جداپذیر باشند، نگاشت کرد. از بین نگاشت‌های مرسوم، می‌توان به نگاشت تابع پایه شعاعی^۲، خطی^۳ و چند جمله‌ای^۴ اشاره کرد (برای کسب اطلاعات بیشتر به مرجع (Vapnik 1999) مراجعه کنید). در این پژوهش عملکرد هر یک از این نگاشت‌ها مورد بررسی قرار گرفته و مناسب‌ترین مدل طبقه‌بندی ساخته می‌شود.

۳-۴) تفکیک‌پذیری ویژگی‌های پیشنهادی

همانطور که در مرجع (Sommer and Golz 2002) ذکر شده است، فشردگی و تفکیک‌پذیری ویژگی‌های مورد استفاده در توصیف داده‌ها، مهم‌ترین عامل برای ایجاد یک مدل طبقه‌بندی کننده کارا است. از این‌رو، هر اندازه که ویژگی‌ها فشرده‌تر و تفکیک‌پذیرتر باشند، به همان اندازه، عملیات جداسازی و طبقه‌بندی با دقت و سهولت بیشتری انجام خواهد گرفت. در این بخش سعی داریم تا با استفاده از یک روش تجربی، میزان کارایی ویژگی‌های پیشنهادی را از نظر فشردگی و تفکیک‌پذیری مورد بررسی قرار دهیم. برای رسیدن به این هدف، از شبکه عصبی خود سازمانده، برای نگاشت فضای ۲۴ مولفه‌ای ویژگی‌های استخراجی به فضای دوبعدی استفاده می‌کنیم. با این کار، می‌توان ساختار داده‌های توصیف‌شده توسط ویژگی‌های پیشنهادی را به صورت بصری مشاهده و ارزیابی کرد.

شبکه‌های عصبی خود سازمانده (SOM) (Kohonen, 1995) یکی از شبکه‌های عصبی پرکاربرد است که در زمره روش‌های یادگیری رقابتی قرار می‌گیرد. مرحله آموزش این شبکه عصبی که

1. Supervised classification
2. Radial Basis Function (RBF)
3. Linear (LIN)
4. Polynomial (POL)

نگاشت دوبعدی در SOM برابر با 300×300 بوده و از فاصله اقلیدسی به‌عنوان معیار شباهت، استفاده شده است. نگاشت‌های تولید شده توسط SOM به‌وضوح نشان می‌دهند که اجتماع بردارهای ویژگی استخراج‌شده از دو فضای رنگی RGB و HSV، به‌طور مطلوب‌تری خوشه‌بندی شده و فشردگی و جدایی‌پذیری آنها به‌طور قابل ملاحظه‌ای، بهتر از حالت‌هایی است که ویژگی‌های آماری دو فضای رنگ، به‌طور مجزا استفاده شده‌اند. همچنین نکته مهم دیگری هم قابل نتیجه‌گیری است و آن این‌که ویژگی‌های استخراج‌شده از هر یک از گروه‌ها (منابع طبیعی و سازه‌های ساخت دست بشر) همگن بوده و شباهت‌های درون‌گروهی زیادی دارند و به همین دلیل فشردگی هر یک از گروه‌ها بسیار مطلوب است. این فشردگی و جدایی‌پذیری، تاثیر به‌سزایی در کارایی مدل طبقه‌بندی‌کننده، خواهد داشت.

۴) آزمایش‌ها

در این بخش، آزمایش‌های طراحی‌شده برای ارزیابی کارایی راهکار پیشنهادی ارائه می‌شوند. ابتدا، جزئیات مربوط به نحوه ایجاد بانک تصاویر و برچسب‌گذاری آنها بیان می‌شود. سپس نتایج آزمایش‌های گسترده‌ای که برای ارزیابی عملکرد راهکار پیشنهادی و عناصر مختلف آن انجام گرفته، مورد بحث و بررسی قرار می‌گیرند، و نهایتاً کارایی راهکار پیشنهادی در مقایسه با سایر روش‌های موجود بررسی و مورد مطالعه قرار می‌گیرد.

به‌صورت بدون ناظر است، باعث شده تا این شبکه عصبی، روش مناسبی برای خوشه‌بندی داده‌هایی باشد که از قبل، هیچ‌گونه اطلاعاتی از خصوصیات آنها در دست نیست. این شبکه عصبی با استفاده از داده‌های در دسترس و الگوریتم یادگیری بدون ناظر، فضای ویژگی‌های توصیف‌کننده داده‌ها را از یک فضای موجود به فضای دوبعدی و یا سه‌بعدی نگاشت می‌کند.

به منظور ارزیابی فشردگی و جدایی‌پذیری ویژگی‌های پیشنهادی بوسیله SOM، ۵۰۰۰ نمونه (قطعه) شامل سازه‌های ساخت بشر و ۵۰۰۰ نمونه شامل منابع طبیعی را به‌صورت تصادفی از میان تصاویر آموزشی انتخاب می‌کنیم (رجوع شود به بخش ۴-۱-الف). اندازه قطعات نمونه‌های انتخاب‌شده در این آزمایش 8×8 پیکسل است. به‌منظور ارزیابی ترکیب‌های مختلف ویژگی‌های آماری استخراج‌شده از فضاهای رنگی RGB و HSV، شبکه‌ی SOM را به‌صورت مجزا یک‌بار با ۱۲ ویژگی استخراج‌شده از فضای رنگی RGB، یک‌بار با ۱۲ ویژگی استخراج‌شده از فضای رنگی HSV و بار دیگر با اجتماع همه ۲۴ ویژگی استخراج‌شده از دو فضای رنگی RGB و HSV آموزش می‌دهیم. نتایج نگاشت‌های دوبعدی در شکل ۳، نشان داده شده است. در این شکل، ناحیه‌های سبز رنگ، نگاشت ویژگی‌های استخراج‌شده از مناطق طبیعی و ناحیه‌های سفید رنگ، نگاشت ویژگی‌های استخراج‌شده از مناطق شامل سازه‌های دست بشر است. توضیح این‌که ابعاد فضای



(ج)



(ب)



(الف)

شکل ۳. نگاشت بردارهای ویژگی استخراج‌شده از (الف) فضای رنگی RGB، (ب) فضای رنگی HSV، (پ) مجموع فضاهای رنگی RGB و HSV. ناحیه‌های سبز رنگ نگاشت ویژگی‌های استخراج‌شده از مناطق طبیعی و ناحیه‌های سفید رنگ نگاشت ویژگی‌های استخراج‌شده از مناطق شامل سازه‌های دست بشر است.

۴-۱) بانک تصاویر

برای ارزیابی راهکار پیشنهادی از دو بانک تصاویر هوایی، استفاده شده است. بانک تصاویر اول که به منظور انجام این پژوهش و بر اساس اهدافی که در آن دنبال می‌شود، ساخته شده است و بانک دوم، شامل تصاویر استاندارد است که در سایر مقالات ارایه شده اخیر، مورد استفاده قرار گرفته است. در ادامه توضیحات لازم درباره این دو بانک ارایه می‌شود.

الف) بانک تصاویر ساخته شده در این پژوهش:

برای ایجاد این بانک تصاویر، تعداد ۱۵ تصویر هوایی با جستجوی عبارت "aerial images" در وبسایت گوگل انتخاب شده‌اند. اندازه تصاویر جمع‌آوری شده، متفاوت بوده و حداقل و حداکثر عرض و یا طول در آنها به ترتیب برابر با ۷۰۰ و ۲۴۰۰ پیکسل است. با توجه به سبب نسبتاً بزرگ تصاویر جمع‌آوری شده، می‌توان به

تعداد زیاد و کافی، قطعه تصویر از آنها به‌عنوان نمونه‌های آموزشی و آزمایشی تهیه کرد. تصاویری که عرض و یا طول آنها، توان صحیحی از ۲ نبودند، به‌صورت دستی بطوری‌که اندازه آن در هر دو بُعد، توانی از ۲ شود، بریده شدند. در شکل ۴، مجموعه تصاویر گردآوری شده نشان داده شده‌اند. توضیح این‌که بانک تصاویر جمع‌آوری شده و برچسب‌گذاری شده در این پژوهش برای انجام کارهای پژوهشی سایر پژوهشگران قابل ارایه و دسترس‌پذیر است.

برای برچسب‌گذاری تصاویر جمع‌آوری شده، پیکسل‌های شامل سازه‌های دست بشر، توسط یک ناظر انسانی مشخص و علامت‌گذاری می‌شوند (در این مقاله با رنگ قرمز^۱). تصاویر علامت‌گذاری شده با چهار اندازه قطعه ۶۴×۶۴، ۳۲×۳۲، ۱۶×۱۶ و ۸×۸ بلوک‌بندی شده و در چهار مجموعه جداگانه قرار می‌گیرند. سپس بر



شکل ۴. تعداد ۱۵ تصویر جمع‌آوری شده از اینترنت با جستجوی عبارت "aerial images" در وبسایت گوگل

۱. R=255, G = 0, B = 0

(ب) بانک تصاویر ماساچوست: این بانک، مربوط به تصاویر مناطق شهری و حاشیه‌ای شهر ماساچوست آمریکا است که در برگیرنده ساختمان‌هایی با اندازه‌های متفاوت بوده و توسط نویسندگان مرجع (Mnih 2013) ایجاد شده است. این مجموعه شامل ۱۵۱ تصویر به ابعاد ۱۵۰۰×۱۵۰۰ پیکسل است که از میان آنها ۱۳۷ تصویر برای آموزش، ۱۰ تصویر برای آزمایش و ۴ تصویر برای ارزیابی زمان آموزش بر اساس سند ارائه شده در (Mnih 2013)، مورد استفاده قرار گرفته‌اند. جزئیات بیشتر در مورد این بانک، در مرجع مذکور ارائه شده است و برخی موارد لازم درباره نحوه انجام آزمایش‌ها و مقایسه با سایر روش‌ها، در قسمت مربوطه (بخش ۴-۶) آورده شده است.

۴-۲) معیارهای ارزیابی

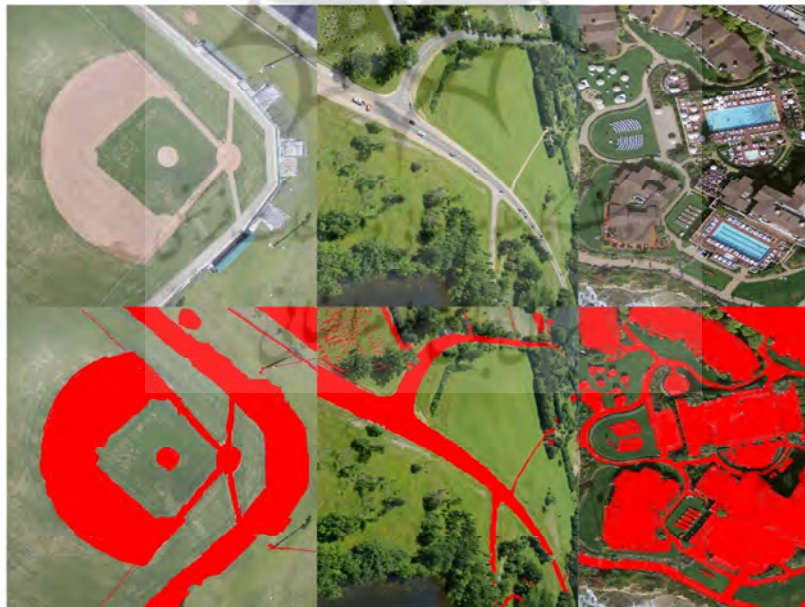
به منظور بررسی کارایی راهکار پیشنهادی از سه معیار صحت (A) ، بازخوانی (R) و دقت (P) که به ترتیب در روابط ۶ الی ۸ تعریف شده‌اند، استفاده می‌کنیم.

اساس رابطه (۵)، قطعات هر مجموعه به نمونه‌های «سازه دست بشر» $(+1)$ و نمونه‌های «منابع طبیعی» (-1) برچسب‌گذاری (L) می‌شوند. در شکل ۵، نمونه‌ای از تصاویر علامت‌گذاری شده نشان داده شده است.

$$L = \begin{cases} +1 & \alpha \leq \theta \leq \beta \\ -1 & \text{Otherwise} \end{cases} \quad (5)$$

در رابطه (۵)، θ ، α و β به ترتیب بیانگر درصد پیکسل‌های قرمز، درصد آستانه‌ی پایین و درصد آستانه‌ی بالا هستند.

با در نظر گرفتن اندازه قطعات تعیین شده، تعداد ۱۱۴۳۳ قطعه تصویر ۶۴×۶۴ پیکسل، تعداد ۴۵۷۳۲ قطعه تصویر ۳۲×۳۲ پیکسل، تعداد ۱۸۲۹۲۸ قطعه تصویر ۱۶×۱۶ پیکسل و تعداد ۷۳۱۷۱۲ قطعه تصویر ۸×۸ از تصاویر جمع‌آوری شده، به دست می‌آیند. از هر یک از مجموعه‌ها، ۶۰ درصد قطعات تصاویر به‌عنوان قطعات آموزشی و ۴۰ درصد آن به‌عنوان قطعات آزمایشی مورد استفاده قرار می‌گیرند. خلاصه داده‌های آموزشی و آزمایشی در جدول ۱ آورده شده است.



شکل ۵. نمونه‌ای از تصاویر علامت‌گذاری شده توسط ناظر انسانی: در تصاویر سطر دوم، نواحی که با رنگ قرمز علامت‌گذاری شده‌اند، نشان‌دهنده‌ی نواحی شامل سازه‌های دست بشر در تصاویر سطر اول است.

1. Accuracy
2. Recall
3. Precision

مطرح شده در رابطه (۵) با مقادیر $\alpha=71\%$ و $\beta=80\%$ در نظر گرفته شده است (رجوع شود به بخش ۴-۴). سیستم پیشنهادی با استفاده از نمونه‌های آموزشی هر یک از مجموعه‌های چهارگانه قطعات، آموزش داده شده و توسط نمونه‌های آزمایشی مربوط به هر مجموعه مورد آزمایش قرار می‌گیرد. جداول ۲، ۳ و ۴ نتایج به‌دست آمده را به ترتیب برای توابع نگاشت LIN، POL و RBF نشان می‌دهند. در نتایج نشان داده شده، بهترین نتایج به صورت پررنگ مشخص شده‌اند. این جداول به وضوح نشان می‌دهند که انتخاب اندازه قطعات برابر با 8×8 ، کارایی بهتری نسبت به سایر اندازه‌ها دارد. با توجه به ریز بودن اشیا ساخت بشر در اغلب تصاویر هوایی، احتمال این که کل یک قطعه حاوی شی مصنوعی باشد، در قطعات کوچک‌تر بیشتر است. از این‌رو، اندازه قطعات 8×8 منجر به نتایج بهتری در مقایسه با اندازه‌های بزرگ‌تر شده است.

شکل ۶، مقایسه کارایی توابع نگاشت استفاده شده در الگوریتم یادگیری SVM را برای حالتی که اندازه قطعات برابر با 8×8 است نشان می‌دهد. با توجه به این شکل، می‌توان به این نتیجه رسید که از نظر کارایی، جایگزینی تابع نگاشت LIN با تابع نگاشت RBF، افزایش جزئی در کارایی راهکار پیشنهادی دارد. همچنین، بررسی‌های عینی نتایج بدست آمده توسط ناظر انسانی نیز نشان می‌دهد که تفاوت قابل ملاحظه‌ای از نظر معنایی بین نتایج بدست آمده توسط توابع نگاشت LIN و RBF وجود ندارد. دلیل این امر را می‌توان در تفکیک‌پذیری خطی مناسب ویژگی‌های استخراج شده دانست. با توجه به اختلاف بسیار جزئی موجود، نویسندگان استفاده از تابع نگاشت LIN را به این دلیل که هم از نظر

$$A = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

$$R = \frac{TP}{TP+FN} \quad (7)$$

$$P = \frac{TP}{TP+FP} \quad (8)$$

در روابط ۶ الی ۸ داریم:

TP: قطعه مورد نظر شامل سازه ساخت بشر است و توسط راهکار پیشنهادی به درستی تشخیص داده شده است.

TN: قطعه مورد نظر شامل منابع طبیعی است و توسط راهکار پیشنهادی به درستی تشخیص داده شده است.

FP: قطعه مورد نظر شامل منابع طبیعی است و توسط راهکار پیشنهادی به اشتباه سازه دست بشر تشخیص داده شده است.

FN: قطعه مورد نظر شامل سازه ساخت بشر است و توسط راهکار پیشنهادی به اشتباه منابع طبیعی تشخیص داده شده است.

۳-۴ نتایج

در این بخش نتایج بدست آمده از آزمایش‌های مختلف ارائه می‌شود. در این پژوهش، از سه تابع نگاشت چند جمله‌ای (POL)، خطی (LIN) و تابع پایه شعاعی (RBF) در الگوریتم آموزشی SVM برای ساخت مدل طبقه‌بندی استفاده شده است. توضیح اینکه، از کتابخانه LibSVM (Chang, Lin et al. 2011) و مقادیر پارامترهای پیش فرض ذکر شده در این کتابخانه برای ساخت مدل مبتنی بر الگوریتم SVM استفاده شده است. همچنین در هر سه آزمایش پارامترهای α و β

جدول ۱. اطلاعات مربوط به تعداد داده‌های آموزشی و آزمایشی چهار مجموعه قطعات بلوک‌بندی شده

اندازه بلوک‌بندی	تعداد کل نمونه‌ها	تعداد نمونه‌های آموزشی	تعداد نمونه‌های آزمایشی	تعداد نمونه‌های آزمایشی مثبت	تعداد نمونه‌های آزمایشی منفی
64×64	۱۱۴۳۳	۶۸۶۰	۴۵۷۳	۱۴۵۶	۳۱۱۷
32×32	۴۵۷۳۲	۲۷۴۳۹	۱۸۲۹۳	۶۴۵۶	۱۱۸۳۷
16×16	۱۸۲۹۲۸	۱۰۹۷۴۷	۷۳۱۷۱	۲۱۹۸۶	۵۱۱۸۵
8×8	۷۳۱۷۱۲	۴۳۹۰۲۷	۲۹۲۶۸۵	۱۳۰۸۸۷	۱۶۱۷۹۸

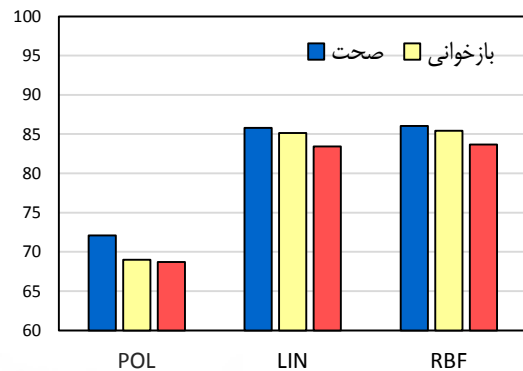
رابطه ۵) در کارایی راهکار پیشنهادی، بازه‌های مختلفی را برای این دو پارامتر آزمایش می‌کنیم. در این آزمایش مقادیر α و β را به ترتیب برابر با ۵۱، ۶۰، ۶۱، ۷۰ و ۷۱ و ۸۰، ۸۱، ۹۰ و ۹۱ در نظر گرفته و کارایی راهکار پیشنهادی را برای هر یک از حالات ارزیابی می‌کنیم. با توجه به نتایج آزمایش‌های قبلی، مدل مبتنی بر SVM (LIN) و اندازه قطعه 8×8 در این آزمایش انتخاب شده است. شکل ۷، نتایج به‌دست آمده را نشان می‌دهد.

بر اساس نتایج نشان داده شده در شکل ۷، می‌توان به این نتیجه رسید که با مقادیر کوچک‌تر آستانه، راهکار پیشنهادی از صحت و دقت پائین‌تری برخوردار است. با افزایش مقادیر آستانه پایین و بالا، عملکرد سیستم بهتر می‌شود. البته همانطور که در شکل مشاهده می‌شود، با مقادیر آستانه‌ی بزرگ‌تر، میزان حساسیت راهکار پیشنهادی افزایش پیدا کرده و در نتیجه کارایی سیستم از نظر دقت کاهش پیدا می‌کند. در مقابل، میزان بازخوانی و صحت سیستم رفته‌رفته افزایش پیدا می‌کند و مجدداً پس از بازه ۷۱-۸۰ شروع به کاهش می‌کند. لذا انتخاب حد آستانه بالا و پایین به ترتیب برابر با ۷۱ و ۸۰ می‌تواند انتخاب مناسبی در راهکار پیشنهادی باشد. نمونه‌ای از نتایج بدست آمده توسط روش پیشنهادی در شکل ۸، نشان داده شده است.

۴-۵) مقایسه با سایر روش‌های استخراج ویژگی
در این قسمت، راهکار پیشنهادی استخراج ویژگی‌های بافت در تصاویر هوایی را با دو روش توصیف بافت مشهور و رایج به‌نام‌های الگوهای دودویی محلی (LBP) (Ojala, Pietikainen et al. 2002) و الگوهای سه‌تایی محلی (LTP) (Tan and Triggs 2010) مقایسه می‌کنیم. الگوی دودویی محلی، روشی شناخته شده برای توصیف بافت بوده و کارایی بسیار خوبی در شرایطی که تغییرات شدت نور وجود داشته باشد، از خود نشان داده است. همچنین از نظر محاسباتی نیز روش بسیار کارآمدی به حساب می‌آید.

- | | |
|--------------------------|---------------------------|
| 1. Off-line | 2. On-line |
| 3. Local binary patterns | 4. Local ternary patterns |

سرعت آموزش (که بصورت برون خط^۱ انجام می‌گیرد) برای ایجاد مدل و هم از نظر سرعت بازناسایی (که بصورت برخط^۲ انجام می‌گیرد) که به مراتب سریع‌تر از تابع نگاشت RBF است، پیشنهاد می‌کنند.



شکل ۶. مقایسه کارایی (بر اساس درصد) توابع نگاشت استفاده شده در الگوریتم یادگیری SVM

جدول ۲. کارایی راهکار پیشنهادی با استفاده از مدل SVM (POL)

اندازه قطعات				
8×8	16×16	32×32	64×64	
۷۲/۰۹	۷۳/۹۴	۶۸/۲۴	۵۶/۴۴	A
۶۹/۰۱	۶۳/۶۴	۶۷/۸۶	۶۰/۴۴	R
۶۸/۷۱	۵۵/۸۲	۵۳/۹۸	۳۸/۳۳	P

جدول ۳. کارایی راهکار پیشنهادی با استفاده از مدل SVM (LIN)

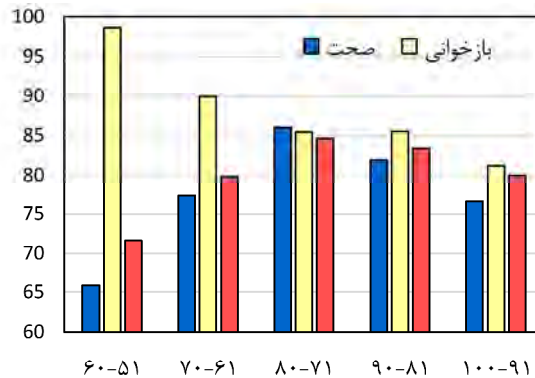
اندازه قطعات				
8×8	16×16	32×32	64×64	
۸۵/۷۹	۸۱/۸۷	۷۳/۷۱	۶۳/۴۸	A
۸۵/۱۳	۷۱/۸۲	۶۷/۸۶	۶۱/۹۵	R
۸۳/۴۲	۶۹/۰۶	۶۱/۵۷	۴۴/۷۰	P

جدول ۴. کارایی راهکار پیشنهادی با استفاده از مدل SVM (RBF)

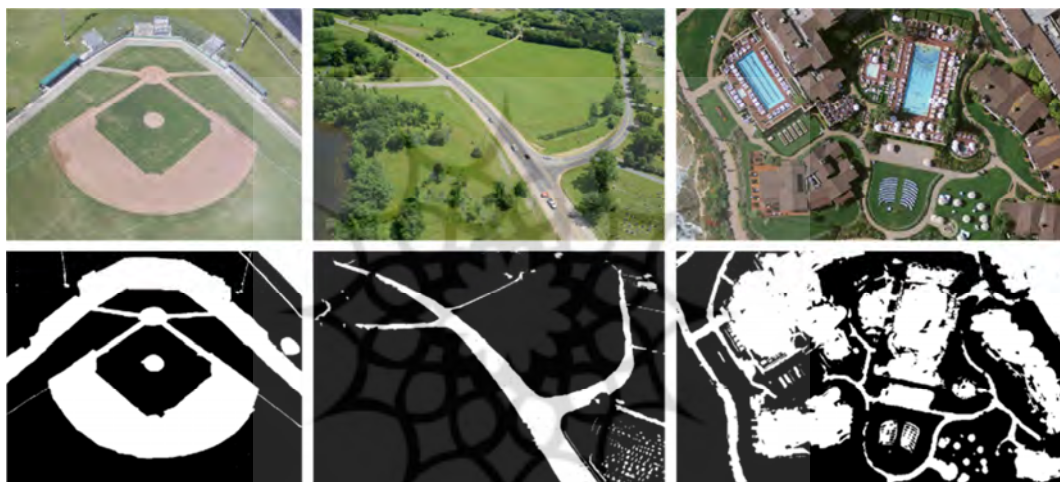
اندازه قطعات				
8×8	16×16	32×32	64×64	
۸۶/۰۳	۸۲/۱۴	۷۷/۵۳	۶۷/۹۰	A
۸۵/۴۴	۷۲/۷۳	۷۷/۱۵	۶۷/۵۸	R
۸۳/۶۶	۶۹/۳۳	۶۵/۴۰	۴۹/۷۰	P

۴-۴) بررسی تاثیر مقدار حد آستانه

برای بررسی تاثیر مقادیر آستانه پایین و بالا (α و β) در



شکل ۷. مقایسه عملکرد (بر اساس درصد) راهکار پیشنهادی با بازه‌های مختلف برای مقادیر آستانه‌ی پایین و بالا در رابطه (۵).

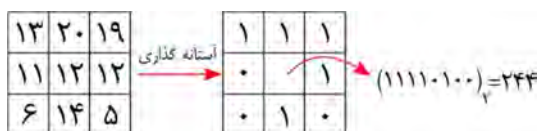


شکل ۸. نمونه‌ای از عملکرد روش پیشنهادی در تشخیص سازه‌های ساخت بشر. سطر اول تصاویر هوایی و سطر دوم تصاویر برچسب‌دار شده متناظر توسط روش پیشنهادی را نشان می‌دهند.

در رابطه (۹) داریم:

$$s(f_n - f_c) = \begin{cases} 1 & f_n \geq f_c \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

در روابط (۹) و (۱۰)، n اندیس پیکسل‌های همسایه، c اندیس پیکسل مرکزی و f_n و f_c به ترتیب میزان شدت روشنایی پیکسل‌های n و c هستند.



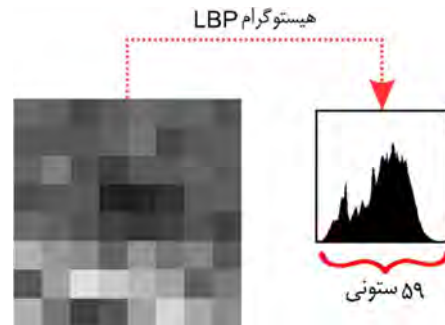
شکل ۱۰. نمونه‌ای از نحوه کار عملگر LBP با همسایگی ۳×۳

عملگر LBP به هریک از پیکسل‌های همسایه پیکسل مرکزی و بر اساس میزان اختلاف هر یک از آنها با پیکسل مرکزی، عدد ۱ یا ۰ اختصاص می‌دهد. سپس عدد دودویی حاصل از کنار هم چیده شدن اعداد ۱ یا ۰ تبدیل به عدد ده‌دهی متناظر شده و پیکسل مرکزی با آن عدد توصیف می‌شود. نحوه کار این عملگر در روابط ۹ و ۱۰ توصیف و نمونه‌ای از انجام آن در شکل ۱۰ نمایش داده شده است. در انتها، تصویر مورد نظر بوسیله همسیتوگرام، مقادیر جدید پیکسل‌ها بازنمایی می‌شود (شکل ۱۱).

$$LBP(i_c, j_c) = \sum_{n=0}^7 2^n s(f_n - f_c), \quad (9)$$

مجموعه ویژگی‌های مختلف نشان می‌دهد. نتایج ارزیابی شده در این جدول به‌وضوح نشان از کارایی بهتر راهکار پیشنهادی در مقایسه با دو روش شناخته شده توصیف بافت دارد. راهکار پیشنهادی به‌ترتیب با افزایش ۴/۶٪ و ۱/۹۲٪ در صحت و ۵/۵۷٪ و ۷/۹۵٪ در بازخوانی، کارایی روش‌های LBP و LTP را افزایش داده است. اما نکته قابل توجه در نتایج بدست آمده این است که کارایی روش‌های LBP و LTP از نظر معیار دقت، تقریباً با راهکار پیشنهادی یکسان است. بنا به تعریف معیار دقت که بیان می‌کند «چه تعداد از قطعاتی که به‌عنوان قطعه شامل سازه ساخت بشر برچسب گذاری شده‌اند، درست هستند»، می‌توان به این نتیجه رسید که روش‌های LBP و LTP در مقایسه با راهکار پیشنهادی از حساسیت بالاتری برخوردار بوده، به‌همین دلیل در فرآیند برچسب‌گذاری وسواس بیشتری به خرج می‌دهند. در تایید این مطلب، می‌توان به پایین بودن میزان بازخوانی این روش‌ها اشاره کرد. از این‌رو، راهکار پیشنهادی نه تنها میزان حساسیت در بازشناسایی را به‌طور قابل توجهی تعدیل کرده، بلکه به دلیل توصیف مناسبی که از بافت تصاویر دارد، کارایی بهتری هم در بازخوانی و هم در صحت از خود نشان داده است.

در جدول ۵، همچنین زمان کل سپری شده (T) بر حسب ثانیه برای طبقه‌بندی همه نمونه‌های آزمایشی (۲۹۲۶۸۵ قطعه تصویر) توسط روش‌های مختلف گزارش شده است. همانطور که این جدول نشان می‌دهد، راهکار پیشنهادی از نظر زمان اجرایی نیز بهتر از روش‌های LBP و LTP عمل کرده و به‌ترتیب ۷/۶ ثانیه و ۱۲/۲ ثانیه زمان اجرایی را کاهش داده است. دلیل عمده دست‌یابی به چنین سرعت پردازشی، اندازه طول بردار ویژگی‌های راهکار پیشنهادی (بطول ۲۴) است که اندازه آن در مقایسه با روش‌های LBP (بطول ۵۹) و LTP (بطول ۱۱۸) به‌ترتیب تقریباً ۶۰٪ و ۸۰٪ کوچکتر است.



شکل ۱۱: هیستوگرام ۵۹ ستونی LBP از یک تصویر ۸×۸

به‌دلیل این‌که روش LBP بر اساس آستانه‌گذاری روی شدت روشنایی پیکسل‌های همسایه استوار است، کارایی آن زمانی‌که تصویر نویزدار و یا تغییرات شدت نور در بعضی از نواحی نامحسوس باشد، کاهش می‌یابد. برای فایق آمدن به چنین مشکلی، تان و همکارش (Tan and Triggs 2010) روش LBP را توسعه داده و با تغییر رابطه (۱۰) به رابطه (۱۱)، روش جدیدی بنام الگوهای سه‌تایی محلی را معرفی کردند.

$$\hat{s}(f_n - f_c, t) = \begin{cases} +1 & f_n \geq f_c + t \\ 0 & |f_n - f_c| < t \\ -1 & f_n \leq f_c - t \end{cases} \quad (11)$$

در رابطه (۱۱)، t حد آستانه‌ای است که توسط کاربر تعریف می‌شود. بر اساس آزمایش‌های انجام شده در مرجع (Tan and Triggs 2010)، $t=5$ پیشنهاد شده است.

LTP را می‌توان با در نظر گرفتن المانهای مثبت و منفی (یعنی +۱ها و -۱ها)، به‌وسیله دو کانال LBP نمایش داد که در این‌صورت هیستوگرام مربوط به هر یک از کانال‌ها بطور مجزا محاسبه شده و در نهایت باهم ادغام می‌شوند. در این مقاله، عملگر $LBP_{8,2}^{u,2}$ (هیستوگرام ۵۹ ستونی، با شعاع همسایگی برابر با ۲ و تعداد همسایه‌های برابر با ۸) روی قطعات ۸×۸ اعمال می‌شود. در نتیجه طول بردار ویژگی در روش استخراج LBP برابر با ۵۹ و در روش استخراج LTP برابر با ۱۱۸ (۲×۵۹) خواهد بود.

جدول ۵، کارایی راهکار پیشنهادی را با استفاده از

قرار دارند. بر همین اساس و جهت مقایسه عادلانه، ما نیز از این سیاست ارزیابی پیروی کرده و طبق روش مذکور، مقدار ρ را برابر با سه قرار می‌دهیم. نکته مهم دیگر، استفاده نویسندگان مقالات مذکور از «نقطه تلاقی منحنی‌های بازخوانی-دقت»^۱ در نمودار ROC^۲ است. بر همین اساس، ما نیز در نتایج ارائه شده برای مقایسه از این معیار استفاده کرده‌ایم. نتایج مقایسه راهکار پیشنهادی با سایر روش‌ها در جدول ۶ آورده شده است.

همانطور که جدول ۶ نشان می‌دهد، راهکار پیشنهادی علی‌رغم استفاده از ویژگی‌های بسیار ساده آماری و تکنیک پایه برای عمل طبقه‌بندی، عملکرد قابل مقایسه و درصد دقت تشخیص نزدیکی با روش‌های مورد مقایسه، که همگی مبتنی بر یادگیری عمیق هستند، از خود نشان داده است. البته با مقایسه میانگین زمان پردازش روش‌های مختلف در جدول ۶، مشخص است که راهکار پیشنهادی از نظر زمان پردازش، عملکرد بسیار خوبی دارد، بطوری‌که مدت زمان پردازش را نسبت به سایر روش‌ها، تقریباً پنج برابر کاهش داده است.^۳ با در نظر گرفتن این نکته که کاهش زمان پردازش در پردازش انبوهی از تصاویر هوایی، که نمونه‌ای از داده‌های کلان^۴ به‌شمار می‌روند، در کاربردهای واقعی^۵ بسیار ارزشمند شمرده می‌شود، می‌توان چنین نتیجه‌گرفت که استفاده از روش‌های ساده نظیر راهکار پیشنهادی در تشخیص سازه‌های

جدول ۵. مقایسه کارایی و سرعت پردازش راهکار پیشنهادی با

روش‌های LTP و LBP		راهکار پیشنهادی	
LTP	LBP		
۸۳/۸۷٪	۸۱/۱۹٪	۸۵/۷۹٪	A
۷۷/۱۸٪	۷۹/۵۶٪	۸۵/۱۳٪	R
۸۲/۱۴٪	۸۳/۰۸٪	۸۳/۴۲٪	P
۳۱/۴ ثانیه	۲۶/۸ ثانیه	۱۹/۲ ثانیه	T

۴-۶) مقایسه با سایر روش‌های تشخیص

در این بخش، راهکار پیشنهادی با چند روش اخیر (Mnih 2013, Saito and Aoki 2015, Saito, Yamashita et al. 2016, Hui, Du et al. 2018) در زمینه تشخیص سازه‌های ساخت دست بشر در تصاویر هوایی مورد مقایسه قرار می‌گیرد. به سه دلیل عمده، این روش‌ها برای مقایسه انتخاب شده است: (۱) استفاده آنها از یادگیری عمیق که به‌عنوان یک روش بسیار کارآمد در هوش مصنوعی در حال فراگیر شدن است، (۲) در دسترس عموم بودن بانک تصاویر استفاده شده در آنها و (۳) فراهم بودن اطلاعات مربوط به برچسب‌گذاری دستی پیکسل‌های تصاویر در بانک تصاویر استفاده شده.

در مراجع مذکور، پژوهشگران با استفاده از بلوک‌بندی تصاویر به قطعات 64×64 پیکسل و اطلاعات از پیش استخراج شده مربوط به برچسب هر یک از پیکسل‌ها و یادگیری عمیق، اقدام به ساخت مدل طبقه‌بندی کننده کرده‌اند. ذکر دو نکته مهم، در انجام این آزمایش و مقایسه نتایج، ضروری است. اول اینکه، در مراجع مذکور استفاده از بازه‌ای بنام ρ در محاسبه میزان بازخوانی و دقت پیشنهاد شده است. در این نحوه ارزیابی، بازخوانی عبارت است از کسری از پیکسل‌های واقعی مربوط به سازه‌های ساخت دست بشر که در بازه ρ پیکسل‌های تشخیص داده شده توسط راهکار پیشنهادی قرار دارند، و دقت عبارت است از کسری از پیکسل‌های تشخیص داده شده توسط راهکار پیشنهادی به‌عنوان سازه ساخت بشر که در بازه ρ از پیکسل‌های واقعی مربوط به سازه‌های ساخت بشر

1. Precision-recall breakeven point

2. Receiver operating characteristic

۳. کدهای الگوریتم‌های مورد مقایسه در مکاتبه با نویسندگان مقالات مربوطه، در اختیار پژوهشگران این مقاله قرار گرفته است و نتایج گزارش شده براساس اجرای همه‌ی الگوریتم‌ها روی یک سخت‌افزار یکسان است. شایان ذکر است که کد مربوط به مقاله (Saito and Aoki 2015) بنا به اظهار نویسندگان، غیر قابل انتشار است. از اینرو، صرفاً مقدار دقت مقاله مذکور براساس گزارش خود مقاله در جدول ۶ آورده شده است.

4. Big data

5. Real applications

همچنین کاری زمان‌بر است. در نتیجه مطالعات زیادی در این راستا صورت گرفته تا با بهره‌گیری از روش‌های مبتنی بر هوش مصنوعی، این فرآیند را بطور خودکار و هوشمند انجام دهند.

در این مقاله، با استخراج اطلاعات آماری ساده‌ای مثل میانگین، میانه، مد و واریانس از فضای‌های رنگی RGB و HSV، و استفاده از آنها برای توصیف بافت نواحی مختلف در تصاویر هوایی، راهکاری برای طبقه‌بندی بخش‌های مختلف تصویر به نواحی طبیعی و نواحی شامل سازه‌های ساخت بشر پیشنهاد شده. به منظور ارزیابی عملکرد راهکار پیشنهادی، بانک تصاویر مناسبی از تصاویر هوایی مختلف گردآوری و برچسب‌گذاری شد. آزمایش‌های متنوعی برای ارزیابی راهکار پیشنهادی بر روی بانک تصاویر تهیه شده، صورت گرفت که نتایج به‌دست آمده، نشان از دقت (۸۵/۷۹٪) و سرعت بالای تشخیص راهکار پیشنهادی داشت. همچنین، مقایسه راهکار پیشنهادی با سایر روش‌های اخیر، مبتنی بر یادگیری عمیق روی بانک تصاویر ماساچوست، نشان داد که استفاده از روش‌های ساده نظیر راهکار پیشنهادی برای پردازش‌های کلان، با توجه به محدودیت‌های سخت‌افزاری، در حال حاضر می‌تواند گزینه مناسبی به‌شمار رود؛ دقت و میانگین زمان پردازش بدست آمده روی این بانک تصاویر با استفاده از روش پیشنهادی به ترتیب ۹۰/۰۷٪ و ۰/۹۶ ثانیه است که در مقایسه با سایر روش‌ها، روش پیشنهادی میانگین سرعت پردازش را هشت برابر افزایش داده و تقریباً ۴٪ از دقت کمتری نسبت به بهترین روش مورد مقایسه برخوردار است.

راهکار پیشنهادی از جنبه‌های مختلفی می‌تواند مورد مطالعه و توسعه بیشتر قرار بگیرد. به‌عنوان مثال، می‌توان از ارتقا راهکار پیشنهادی برای کلاس‌بندی انواع کلی اشیاء تشخیص داده‌شده در نواحی مربوط به سازه‌های دست بشر نام برد. تفکیک جاده‌ها از ساختمان‌ها، تشخیص مناطق شهری از روستایی و موارد متعدد دیگری که هر یک می‌توانند نقش مهمی

ساخت بشر در تصاویر هوایی، می‌تواند یکی از گزینه‌های بسیار مناسب محسوب شود. همچنین، با توجه به ماهیت روش‌های مبتنی بر یادگیری عمیق، ممکن است مدت زمان استخراج ویژگی‌ها، آموزش و ساخت مدل یادگیری در آنها روزها و یا هفته‌ها به‌طول بیانجامد که این امر نیز ممکن است در مواردی که سخت‌افزار و زیرساخت‌های لازم مهیا نباشد، باعث عملیاتی نشدن این روش‌ها شود. علاوه بر این، در دسترس بودن حجم زیادی از داده‌های آموزشی نیز جزء ملزومات ساخت مدل‌های مبتنی بر یادگیری عمیق به‌شمار می‌رود و در مواردی که دسترسی به داده‌های آموزشی کافی، مقدور نباشد، چنین روش‌هایی از کارایی قابل قبولی برخوردار نخواهند بود (Mnih 2013).

جدول ۶. مقایسه کارایی و سرعت پردازش راهکار پیشنهادی با سایر روش‌ها

میانگین زمان پردازش (ثانیه)	نقطه تلاقی محنی‌های بازخوانی-دقت (%)	راهکار پیشنهادی (Mnih 2013)
۰/۹۶	۹۰/۰۷	Saito, Yamashita et al. (2016)
۴/۸۸	۹۲/۱۱	(Saito and Aoki 2015)
۵/۱۲	۹۴/۲۶	(Hui, Du et al. 2018)
-	۹۲/۳۰	
۷/۸۷	۹۴/۲۳	

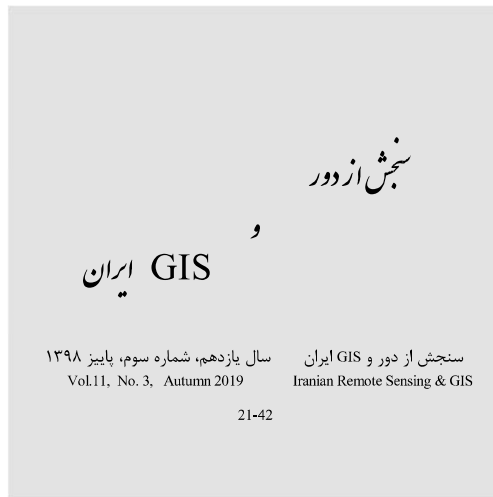
۵) نتیجه‌گیری و محورهای مطالعه و توسعه بیشتر
 رشد روز افزون جمعیت و توسعه جوامع بشری سبب شده تا اطلاعات مکانی جمع‌آوری شده از تصاویر هوایی، یکی از مهمترین ابزارها در کاربردهایی نظیر برنامه‌ریزی شهری، مدیریت بحران، ایجاد و به‌روز رسانی پایگاه داده‌های موجود، شناسایی تغییرات و ... به‌شمار آید. در بسیاری از این کاربردها، لازم است اطلاعات دقیق و به‌روزی از نواحی ساخت بشر در مناطق شهری و نیمه‌شهری در دسترس باشد. تهیه و برچسب‌دار کردن چنین مناطقی، بر روی انبوهی از تصاویر هوایی و به‌روز رسانی آنها، بسیار پرهزینه بوده و

- Chang, C.-C., C.-J. J. A. t. o. i. s. Lin and technology, 2011, **LIBSVM: A library for support vector machines**, 2(3): 27.
- Cheng, G. & J. Han, 2016, **A survey on object detection in optical remote sensing images**, ISPRS Journal of Photogrammetry and Remote Sensing 117: 11-28.
- Cheng, G., J. J. I. J. o. P. Han & R. Sensing, 2016, **A survey on object detection in optical remote sensing images**, 117: 11-28.
- Cheriyadat, A. M. J. I. T. o. G. & R. Sensing, 2014, **Unsupervised feature learning for aerial scene classification**, 52(1): 439-451.
- Cunningham, J. P. & Z. J. T. J. o. M. L. R. Ghahramani, 2015, **Linear dimensionality reduction: Survey, insights, and generalizations**, 16(1): 2859-2900.
- Douglas, S. A. & A. E. J. A. T. o. G. Kirkpatrick, 1999, **Model and representation: the effect of visual feedback on human performance in a color picker interface**, 18(2): 96-127.
- Fauqueur, J., N. Kingsbury & R. Anderson, 2005, **Semantic discriminant mapping for classification and browsing of remote sensing textures and objects**, Image Processing, 2005. ICIP 2005. IEEE International Conference on, IEEE.
- Fischer, A., T. H. Kolbe, F. Lang, A. B. Cremers, W. Förstner, L. Plümer, V. J. C. V. Steinhage & I. Understanding, 1998, **Extracting buildings from aerial images using hierarchical aggregation in 2D and 3D**, 72(2): 185--203.
- Gruen, A., H. J. I. J. o. P. Li & R. Sensing, 1995, **Road extraction from aerial and satellite images by dynamic programming**, 50(4): 11-20.
- Hinz, S., 2003, **Detection and counting of cars in aerial images**, Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on, IEEE.
- Hui, J., M. Du, X. Ye, Q. Qin, J. J. I. G. Sui & R. S. Letters, 2018, **Effective Building Extraction From High-Resolution Remote Sensing Images With Multitask Driven Deep Neural Network**, در کاربردهای گوناگون داشته باشند و به‌عنوان مسیرهای آینده پژوهش، مد نظر قرار گیرند. همچنین بهره‌گیری از سایر انواع ویژگی‌ها نظیر اطلاعات آماری نقاط ویژگی تصاویر به همراه ویژگی‌های معرفی شده در این مقاله، می‌تواند سرآغاز پژوهش‌های بیشتری جهت توسعه سیستم پیشنهادی باشد.
- ۶ منابع
- سجادی، س. ی.، پارسیان، س.، ۱۳۹۶، **استخراج ساختمان به کمک تلفیق داده‌های ابر طیفی و لیدار به روش یادگیری ماشین**، سنجش از دور و GIS ایران جلد دهم، شماره ۲، صص ۱۴-۱.
- Barzohar, M., D. B. J. I. T. o. P. A. Cooper and M. Intelligence 1996, **Automatic finding of main roads in aerial images by using geometric-stochastic models and estimation**, 18(7): 707-721.
- Bhagavathy, S., B. S. J. I. T. o. G. Manjunath and R. Sensing, 2006, **Modeling and detection of geospatial objects using texture motifs** 44(12): 3706-3715.
- Bhagavathy, S., S. Newsam and B. Manjunath, 2002, **Modeling object classes in aerial images using texture motifs**. **Pattern Recognition**, 2002. Proceedings. 16th International Conference on, IEEE.
- Burkov, A., 2019, **The Hundred-Page Machine Learning Book**, Andriy Burkov.
- Burochin, J.-P., B. Vallet, M. Brédif, C. Mallet, T. Brosset, N. J. I. J. o. P. Paparoditis and R. Sensing, 2014, **Detecting blind building façades from highly overlapping wide angle aerial imagery**, 96: 193-209.
- Cai, F., H. Chen and J. J. I. J. I. T. C. S. Ma , 2011, **Man-made object detection based on texture clustering and geometric structure feature extracting**, 3(2): 9-16.
- Cao, G. and X. J. I. J. o. R. S. Yang, 2007, **Man-made object detection in aerial images using multi-stage level set evolution**, 28(8): 1747-1757.

- Hussnain, Z., S. O. Elberink, G. J. I. A. o. t. P. Vosselman, Remote Sensing and S. I. Sciences, 2016, **AUTOMATIC FEATURE DETECTION, DESCRIPTION AND MATCHING FROM MOBILE LASER SCANNING DATA AND AERIAL IMAGERY**, 41.
- Kohonen, T., J. Hynninen, J. Kangas, J. Laaksonen, S. J. U. o. T. L. o. C. PAK & H. Information Science, Finland, 1995, **The self-organizing map program package**.
- Li, X., S.-C. Chen, M.-L. Shyu & B. J. t. I. C. o. D. M. S. Furht, San Francisco Bay, California, USA, 2002, **Image retrieval by color, texture, and spatial information**, 1-8.
- Li, Z. & L. J. I. T. o. I. P. Itti, 2011, **Saliency and gist features for target detection in satellite images**, 20(7): 2017-2029.
- Lin, C., R. J. C. v. Nevatia and i. understanding, 1998, **Building detection and description from a single intensity image**, 72(2): 101-121.
- Ma, W. Y. & B. J. J. o. t. A. S. f. I. S. Manjunath, 1998, **A texture thesaurus for browsing large aerial photographs**, 49(7): 633-648.
- Marcu, A. E. J. a. p. a., 2016, **A local-global approach to semantic segmentation in aerial images**.
- Mnih, V., 2013, **Machine learning for aerial image labeling**, University of Toronto (Canada).
- Moranduzzo, T., F. J. I. T. o. G. Melgani & R. Sensing, 2014, **Automatic car counting method for unmanned aerial vehicle images**, 52(3): 1635-1647.
- Nagarajan, S., T. J. I. J. o. P. Schenk & R. Sensing, 2016, **Feature-based registration of historical aerial images by area minimization**, 116: 15-23.
- Ojala, T., M. Pietikainen, T. J. I. T. o. p. a. Maenpaa and m. intelligence, 2002, **Multiresolution gray-scale and rotation invariant texture classification with local binary patterns**, 24(7): 971-987.
- Ozdemir, B. & S. Aksoy, 2010, **Image classification using subgraph histogram representation**, Pattern Recognition (ICPR), 2010 20th International Conference on, IEEE.
- Penatti, O. A., K. Nogueira & J. A. dos Santos, 2015, **Do deep features generalize from everyday objects to remote sensing and aerial scenes domains?** Proceedings of the IEEE conference on computer vision and pattern recognition workshops.
- Ruiz, L., A. Fdez-Sarria & J. Recio, 2004, **Texture feature extraction for classification of remote sensing data using wavelet decomposition: a comparative study**, 20th ISPRS Congress.
- Saito, S. & Y. Aoki, 2015, **Building and road detection from large aerial imagery**. Image Processing: Machine Vision Applications VIII, International Society for Optics and Photonics.
- Saito, S., T. Yamashita & Y. J. E. I. Aoki, 2016, **Multiple object extraction from aerial imagery with convolutional neural networks**, 2016(10): 1-9.
- Sharma, M. & S. Singh, 2001, **Evaluation of texture methods for image analysis**, Intelligent Information Systems Conference, The Seventh Australian and New Zealand 2001, IEEE.
- Sommer, D. & M. Golz, 2002, **Multiple training of vector-based neural networks to detect density centers in input space**, Proceedings of the European symposium on intelligent technologies, hybrid systems and their implementation on smart adaptive systems.
- Stehling, R. O., M. A. Nascimento & A. X. Falcão, 2002, **A compact and efficient image retrieval approach based on border/interior pixel classification**, Proceedings of the eleventh international conference on Information and knowledge management, ACM.
- Tan, Q., J. Wang & D. A. Aldred, 2008, **Road vehicle detection and classification from very-high-resolution color digital orthoimagery based on object-oriented method**. Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International, IEEE.

- Tan, X. & B. J. I. t. o. i. p. Triggs, 2010, **Enhanced local texture feature sets for face recognition under difficult lighting conditions**, **19**(6): 1635-1650.
- Thuy, N. T. J. D. d., Graz University of Technology, 2009, **Object Detection from Aerial Image**.
- Vapnik, V. N. J. I. t. o. n. n., 1999, **An overview of statistical learning theory**, **10**(5): 988-999.
- Wei, W., Y. J. I. Xin & V. Computing (2010), **Rapid, man-made object morphological segmentation for aerial images using a multi-scaled, geometric image analysis**, **28**(4): 626-633.
- Xia, Y., J. Chen, J. Li & Y. J. M. s. Zhang, 2016, **Geometric discriminative features for aerial image retrieval in social media**, **22**(4): 497-507.



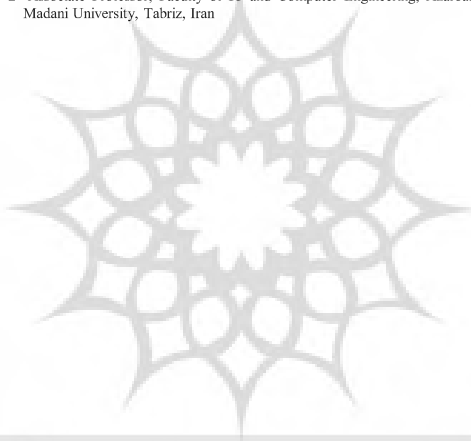


Man-Made Object Detection in Aerial Images Using Color Statistical Features and Machine Learning

Farajzadeh, N.*¹ and Hashemzadeh, M.²

1-Associate Professor, Faculty of IT and Computer Engineering, Azarbaijan Shahid Madani University, Tabriz, Iran

2- Associate Professor, Faculty of IT and Computer Engineering, Azarbaijan Shahid Madani University, Tabriz, Iran



Abstract

Generally, the photos captured by drones and satellites include both natural scenes and man-made objects. Having these two categories classified, we will be able to extract important information from aerial scenes such as the shapes and the alignments of the structures and then, create labeled aerial images accordingly. Obtaining such information is of great interest in, for example, military, urban, and environmental protection applications. However, due to a huge amount of data that is collected in form of images, it seems that manually processing of such data is impossible. Therefore, employing automatic techniques based on artificial intelligence has become more on demand. There are numerous researches on this topic from which detection of buildings, vehicles, roads, and vegetation are of more interest. In this paper, we aim to introduce a method to detect man-made objects in aerial images based on a new set of color statistical features, which can be easily extracted, together with a learning model. Experimental results on a publicly available dataset, Massachusetts dataset, have shown promising results in terms of both accuracy and processing time; the accuracy and the average processing time are 90.07% and 0.96 seconds, respectively.

Keywords: aerial images, natural scene, man-made objects, machine learning, statistical features