

تشخیص وسیله نقلیه درعکس های هوایی با دقت بالا توسط ارائه پراکنده و

سوپرپیکسل ها

چکیده

این مقاله مطالعه ای در مورد تشخیص وسایل نقلیه از نقشه های هوایی با وضوح و دقت بالا را ارائه می دهد. در این مقاله، روش بخش بندی سوپرپیکسل طراحی شده برای عکس های هوایی برای کنترل بخش بندی با نرخ شکست پایین پیشنهاد شده است. برای کارایی بیشتر تمرین و تشخیص، تکه های مهم و عمده ای مبتنی بر مراکز سوپرپیکسل های بخش بندی شده را استخراج کردیم. بعد از بخش بندی، با استفاده استراتژی تکرار انتخاب نمونه های آموزشی که مبتنی بر ارائه پراکنده است، از مجموعه کلی اصلی زیرمجموعه ای آموزشی و کوچک و کامل فراهم می آوریم. توسط زیرمجموعه آموزشی انتخاب شده، فرهنگی دارای توانایی تمییز و تفاوت قائل شدن برای تشخیص وسایل نقلیه فراهم می آوریم. طی آموزش و تشخیص، شبکه مشخصه های گرادیان متمایل هیستوگرام (HOG) به منظور استخراج ویژگی ها استفاده شده اند. برای بهبود بیشتر کارایی تشخیص و آموزش، روشی برای تخمین جهت اصلی مشخص شده برای هر تکه و قطعه ارائه شده است. با چرخش هر قطعه به جهت اصلی آن، جهات ثابتی به قطعات می دهیم. تحلیل های جامع و مقایسات بر روی دو مجموعه داده ها، عملکرد رضایت بخش الگوریتم پیشنهادی را نشان می دهد.

اصطلاحات فهرست - تصویر هوایی، وضوح بالا، نمایش پراکنده، سوپرپیکسل، تشخیص وسیله نقلیه.

1. مقدمه

به دلیل توسعه اقتصادی و نیاز در حال افزایش مسافرت سریع و راحت، اتومبیل ها به شدت محبوب گشته اند . در زندگی روزمره نقش مهمی ایفا می کنند. تعداد زیاد ماشین ها، فشار زیادی بر حمل و نقل، جاده، و مقامات نظارتی ترافیک متحمل می کند و همچنین نظارت ماشین ها را برای جمع آوری اطلاعات ترافیک، راهبندان و پیشگیری ازدحام، کنترل تصادف ترافیک، آمار رفت و آمد وسایل نقلیه، طراحی شبکه جاده، و تخمین حالات پارک حیاتی می سازد.

تعداد زیادی از سنسورهای زمینی ثابت شده مانند حلقه های القایی، سنسورهای پل، دوربین های ثابت، و سنسورهای رادار، برای نظارت و کنترل کارآمد وسایل نقلیه و جمع آوری اطلاعات ترافیکی مورد نیاز هستند. با استفاده از این سنسورهای زمینی ثابت شده، جریان ترافیک، تراکم وسایل نقلیه، و وضعیت پارک تا اندازه ای بدست می آید. با این وجود، این روش ها در ارائه خلاصه کلی از وضعیت ترافیک ناتوان هستند و وضعیت ترافیک منبع اطلاعات حیاتی برای مطالعه نقشه و طرح شبکه جاده ای، مدل سازی، بهینه سازی، و آمار مربوط به ترافیک است.

نیاز به جمع آوری خلاصه کلی وضعیت های ترافیک منجر به نظارت وسایل نقلیه توسط روش های جایگزین مانند تصاویر حسگر از دور گرفته شده توسط ماهواره ها یا هواپیما ها می شود. با توجه به قابلیت آنها برای تهیه پوشش کلی منطقه مورد نظر، تصاویر حسگر از دور بصورت گسترده به منظور نظارت وسایل نقلیه بکار گرفته شده اند. هم اکنون، ماهواره های تجاری پایش زمین وجود دارند مانند IKONOS, GeoEye TWorldView-2, WorldView-3 & QuickBird که تصاویری در دسترس عموم با دقت و وضوح فضایی یک submeter فراهم می کنند. با استفاده از دقت و وضوح فضایی بالا، تصاویر ماهواره ای منبع داده ای برای مطالعه نظارت وسایل نقلیه است. در مقایسه با تصاویر ماهواره ای، معمولاً تصاویر هوایی بدلیل وضوح فضایی زیاد آنها در محدوده 0.1 تا 0.5 متر و بدست آوردن داده آسان تر آنها ترجیح داده می شوند. با وضوح فضایی بیشتر، وسایل نقلیه و حتی ماشین ها در تصاویر هوایی قابل شناسایی هستند. بنابراین، شناسایی وسایل نقلیه از تصاویر هوایی با وضوح بالا برای نظارت ترافیک و کاهش در یک منطقه و بخش بزرگ قابل توجه است. شناسایی دستی و غیر خودکار وسایل نقلیه از تصاویر هوایی زمان بر و کار فشرده ای

است. بنابراین، بوجود آوردن روش اتوماتیک و خودکار شناسایی وسایل نقلیه از تصاویر هوایی با وضوح فضایی بالا امری ضروری و حیاتی است.

برعکس، شناسایی اتوماتیک وسایل نقلیه از تصاویر هوایی با وضوح بالا کار و وظیفه چالش برانگیزی است چرا که وجود ساختارهای متعدد (مانند سطوح های زباله، علائم جاده ای، واحد های الکتریکی، و واحد های تهویه هوا در بالای ساختمان ها) بخصوص در مناطق شهری، می تواند باعث ایجاد هشدارهای خطا شود. علاوه بر این، انسداد های مقطعی که توسط سایه درخت ها و ساختمان ها ایجاد شده ممکن است مشکلات شناسایی وسایل نقلیه را زیاد افزایش دهد. وضعیت روشی عامل حیاتی دیگری در شناسایی وسایل نقلیه از تصاویر هوایی است. نمونه های آموزشی در شناسایی هدف نقش مهمی ایفا می کنند. به منظور بدست آوردن دقت طبقه بندی زیاد مابین وسایل نقلیه و پس زمینه ها، یک مجموعه نمونه آموزشی که شامل انواع مثبت و منفی نمونه ها است مورد نیاز است. آسان ترین راه، استفاده از تمام مجموعه نمونه های آموزشی است. ولی کل مجموعه نمونه آموزشی معمولا خیلی بزرگ و زائد است که باعث پیچیدگی زیاد محاسباتی تشخیص یا آموزش می شود. بنابراین برای آموزش یه عنصر طبقه بندی کننده، انتخاب زیرمجموعه ای کامل و کوچک از نمونه های آموزشی ضروری است. با این وجود انتخاب دستی و غیر خودکار تمامی نمونه های نماینده از تعداد زیادی از منفی ها کار سخت و زمان بری است. علاوه بر این، هر دوی روش های دستی و تصادفی انتخاب نمونه های آموزشی نمیتوانند تضمین کنند که عنصر طبقه بندی کننده ای بهینه را آموزش دهند تا بهترین عملکرد را بدست آورد.

برای بهبود کارایی شناسایی و تشخیص و تشکیل خودکار مجموعه آموزشی نوعی و گویا و کامل، الگوریتم و روشی را توسعه دادیم که از ارائه پراکنده و بخش بندی سوپرپیکسل برای تشخیص اتوماتیک وسایل نقلیه در تصاویر هوایی با وضوح بالا استفاده می کند. برای حرکت دادن پنجره تشخیص بدون اثرات جانبی، بخش بندی مبتنی بر سوپرپیکسل برای بخش بندی کردن تصویر هوایی با وضوح بالا به مجموعه ای از سوپرپیکسل ها ارائه شده است. بر اساس مراکز سوپرپیکسل ها، از این رو قطعات مهم تولید شده اند. سپس ارائه پراکنده برای یادگیری لغت نامه و پردازش طبقه بندی بکار گرفته شده است. برای تشکیل زیرمجموعه آموزشی بهینه، تکرار استراتژی انتخاب نمونه مبتنی بر ارائه پراکنده را

پیشنهاد می دهیم. در طی انتخاب نمونه های آموزشی، کامل بودن نمونه های مثبت و منفی نوعی و گویا در نظر گرفته می شود. با مجموعه آموزشی بهینه انتخاب شده، فرهنگ لغت ارائه پراکنده با قابلیت تمایز زیاد برای تشخیص وسایل نقلیه بدست می آوریم.

روش خود را بر روی 2 مجموعه داده های تصویر هوایی با وضوح بالا اعمال می کنیم. یکی از مجموعه داده ها تصاویر هوایی است که شهر تورنتو کانادا را با وضوح فضایی 0.15 متر پوشش می دهند؛ مجموعه داده دیگر از مجموعه داده تحقیقات بالاسری تصوراتی (OIRDS) است. تحلیل های آزمایشی و مقایسات بر روی هر دو مجموعه داده عملکرد بهتر روش ما در مقابل چندین روش بروز و آخرین فناوری که شامل گرادیان متمایل هیستوگرام (HOG) + ماشین بردار پشتیبان خطی (SVM)، تبدیل ویژگی غیرقابل تغییر با مقیاس (SIFT) + SVM خطی + SVM کرنل و مرکزی.

2. کارهای مرتبط

ارائه پراکنده و بخش بندی سوپرپیکسل در از لحاظ و دید کامپیوتری توجه چشمگیری دریافت کرده اند. ارائه پراکنده بصورت موفقیت آمیزی در زمینه های مختلفی شامل تشخیص صورت، طبقه بندی اهداف و اشیا، طبقه بندی تصویر، مرتفع سازی نويز از تصویر، بازسازی تصویر، برجستگی عینی و فشرده سازی داده ها بکار گرفته شده است. Yokoya و Iwasaki از ارائه پراکنده در تشخیص اهداف و اشیا در تصاویر حسی از راه دور استفاده کردند و نتایج خوبی بدست آوردند. توسعه بخش بندی سوپرپیکسل راه و روشی جدید برای پردازش اولیه تصویر، بخش بندی تصویر، استخراج ویژگی ها، و دنبال کردن اهداف و اشیا فراهم می آورد. در سال های اخیر، بیشتر تحقیقات بر بخش بندی تصویر بر اساس سوپرپیکسل متمرکز شده اند و بسیاری از رویکردها بوجود آمده اند. رویکرد های نوعی و گویا شامل الگوریتم های مبتنی بر نمودار و الگوریتم های مبتنی بر گرادیان می شوند. موفقیت های اخیر خوشه بندی خطی تکراری (SLIC)، ورونی مرکزی وزن شده بر اساس لبه ها مبتنی بر موزاییک (قطعه قطعه) (سلول های V)، و خوشه بندی نرخ

آنتالپی هستند. استفاده از ارائه پراکنده و بخش بندی سوپرپیکسل راهی جدید برای تشخیص وسایل نقلیه در تصاویر هوایی با وضوح بالا است.

رویکردهای بسیاری برای تشخیص وسایل نقلیه در تصاویر هوایی با وضوح بالا بوجود آمده اند. اکثر رویکردها به دو نوع مدل های وسایل نقلیه مانند مدل های ضمنی مبتنی بر ظاهر و مدل های صریح و روشن بخش پذیر هستند.

مدل های ضمنی مبتنی بر ظاهر: یک مدل ضمنی مبتنی بر ظاهر معمولا شامل شدت تصویر یا ویژگی های بافتی می شود که توسط پنجره یا هسته کوچک محاسبه می شوند که پیکسل داده شده یا خوشه ای کوچکی از پیکسل ها را احاطه می کند. سپس، تشخیص توسط آزمایش بردار های ویژگی پیکسل های اطراف میانی تصویر انجام می گیرد. Cheng و همکارانش با استفاده از شبکه های Bayesian پویا برای تشخیص وسایل نقلیه از نظارت هوایی، نتایج امید دهنده ای در مجموعه داده های چالش بر انگیز بدست آوردند. ولی مدل رنگ، که مخصوصا برای جداسازی ماشین ها از زمینه طراحی شده هنوز هم نمی تواند از تشخیص اشتباه دلیل همپوشانی ماشین ها و مدل های رنگ زمینه جلوگیری کند. مشکل دیگر این است که رویکرد باید باعث ترکیب پیکسل های تشخیص داده شده به تک وسایل نقلیه شود که زمانی که وسایل نقلیه در مجاورت کوچکی پارک شده اند، وظیفه ای مشکل و سخت می شود. علاوه بر این، آزمایش و بررسی تشخیص در تمامی پیکسل ها، نه تنها پیچیدگی محاسبات را افزایش می دهد بلکه نرخ تشخیص اشتباه را افزایش می دهد. Shao و همکارانش ابتدا تشخیص وسایل نقلیه با استفاده از ویژگی های متعدد (برای مثال HOG، طرح محلی باینری، و هیستوگرام مخالف) و هسته تقاطع SVM را بررسی کردند. مشابه آن Moranduzzo، Melgani و SIFT و SVM را برای تشخیص ماشین ها از تصاویر وسایل نقلیه هوایی غیرنامگذاری شده ترکیب کردند. Kembhavi و همکارانش ماشین ها را از تصاویر هوایی با وضوح بالا با بکارگیری حداقل مربعات جزیی، تحلیل قدرتمند انتخاب ویژگی، و توضیح دهنده ویژگی اضافی تشخیص دادند که شامل نقشه های احتمال رنگ، ویژگی های HOG، و مقایسات زوج پیکس ها هستند که ویژگی های ساختاری یک ماشین را بدست می آورند. کار آنها عملکردی چشمگیر را نشان می دهد. Moranduzzo و Melgani رویکردی مبتنی بر کاتالوگ برای تشخیص ماشین ها در تصاویر UAV پیشنهاد دادند. ولی کاربرد آن به حس های مخصوص محدود است

چرا که باید از جاده آسفالت شده به عنوان راهنمای اولیه استفاده کند. مشکل دیگر این است که باید پیکسل های تشخیص داده شده را در وسایل نقلیه فرد و تک (منظور تمامی پیکسل هارا ترکیب کرده و بعنوان یک عکس به عنوان وسیله نقلیه نمایش دهد) با ترکیب کند. Hinz و Baumgartner ویژگی های وسایل نقلیه بر اساس مدل وسیله نقلیه سلسله مراتبی استخراج کردند که سطوح مختلفی را شرح می دهد. Khan و همکارانش ویژگی های وسیله نقلیه را بر اساس مدل 3بعدی استخراج کردند. Wang و همکارانش مدل شکل ضمنی و رای دادن Hough را برای تشخیص ماشین در نقاط ابری 3بعدی بکار گرفتند که نتایج چشمگیری حاصل شد.

مدل های صریح و روشن: در مورد مدل صریح و واضح، یک وسیله نقلیه معمولا توسط یک جعبه، ارائه به عنوان قاب، و یا مدل مورفولوژیکی توصیف می شود. تشخیص ماشین توسط تطابق مدل ماشین با تصویر با استفاده از استراتژی "بالا-پایین" یا "پایین-بالا" انجام می گیرد. Zheng و همکارانش از تبدیل گشایش سیاه و سفید و تبدیل بالا-کلاه سیاه و سفید برای شناسایی وسایل نقلیه احتمالی در پس زمینه روشن و یا سفید استفاده کرده و از تبدیل خاتمه دهنده سیاه و سفید و تبدیل پایین-کلاه برای شناسایی وسایل نقلیه احتمالی در پس زمینه های سیاه و بی نور استفاده کردند. سپس اطلاعات سائز و اندازه برای حذف هشدار های اشتباه بکار گرفته می شود. رویکرد آنها عملکرد خوبی در تصاویر هوایی بزرگ راه ها نشان می دهد ولی تخمین های مقدار خاکستری پس زمینه و اطلاعات جغرافیایی داده های سیستم مورد نیاز است. در نتیجه، این روش برای صحنه های عمومی مناسب نیست. یک وسیله نقلیه همچنین به عنوان جعبه 3بعدی با ابعاد عرض و طول و ارتفاع نمایش داده شده است.

مطالعات زیادی انتخاب نمونه از مقادیر زیادی داده های آموزشی را مطالعه کرده اند. Zhou و همکارانش روش کاهش نمونه برای متوجه کردن مشکل غیرمتعادل بودن نمونه های SVM پیشنهاد داد. Nie و همکارانش روش فعالی برای انتخاب نمونه های با بیشترین نمایندگی و نمایان گری (منظور جامع بودن نمونه ها به نمایندگی از بقیه نمونه ها) برای برچسب زدن در مرحله آموزش فعال اولیه برای کاهش کارها و وظایف دستی و غیر خودکار پیشنهاد داد. آنها از روش تکرار برای انتخاب زیرمجموعه های نمونه های با نمایندگی بیشتر با استفاده از نرم ها و استانداردهای موثر و ایجاد کننده پراکندگی استفاده کردند. با این وجود این روش هنوز زمان بر است.

معمولا دو عیب و نقص روش های تشخیص ماشین با استفاده از مدل صریح و روشن مشخص و واضح است. اول، تشخیص معمولا بر اساس لبه های مشخص شده است که منجر به نامستحکم بودن در برابر نویز و پس زمینه پیچیده می شود. دوم، این روش ها عملکرد ضعیفی در حالات انسداد جزئی و تغییرات شکل دارند چرا که مدل های ماشین ها از قبل بصورت شدید و استواری تعریف شده اند. اکثر روش های تشخیص ماشین بروز (حتی نرم افزار های تجاری) برای حس از دور دست تصاویر، از مدل ضمنی بدلیل توانایی تعمیم بهتر آن استفاده می کنند ولی روش های موجود که مدل ضمنی را استفاده می کنند هنوز از 2 مشکل زیر رنج می برند.

اول اکثر روش ها بر اساس پیکسل هستند و یا از یک پنجره لغزنده با مرحله لغزنده حین تشخیص استفاده می کنند. روش های مبتنی بر پیکسل محاسبات سنگینی دارند. علاوه بر این، این روش ها باید پیکسل های تشخیص داده شده را به یک عکس تک باز ترکیب کنند که وقتی وسایل نقلیه در محدوده نزدیک بهم پارک شده باشند امری سخت و مشکل می شود. در روش های پنجره لغزنده، مرحله و گام لغزنده سرعت پردازش و نرخ فراخوانی تشخیص را تحت تاثیر قرار می دهد. یک گام لغزنده بزرگ ممکن است باعث سرعت زیاد پردازش شود ولی باعث کاهش در نرخ فراخوانی می شود. یک گام لغزنده کوچک ممکن است نرخ فراخوانی را کاهش دهد ولی منجر به هزینه محاسباتی زیاد می شود. برقراری تعادل بین نرخ فراخوانی تشخیص و سرعت پردازش مشکل است. استراتژی کارآمدتر اسکن کردن برای بهبود کارایی اسکن کردن مطلوب است.

دوم، نمونه های آموزشی بصورت دستی یا حتی تصادفی انتخاب شده اند. انتخاب دستی نمونه های آموزشی زمان بر است. برای تشخیص وسایل نقلیه در تصاویر هوایی، پس زمینه پیچیده باعث تعداد زیاد نمونه های منفی می شود که انتخاب دستی زیرمجموعه آموزشی منفی بهینه را دشوار می سازد. با در نظر گرفتن روش انتخاب تصادفی، ممکن است باعث ایجاد عملکرد غیر پایدار شده و معمولا به عملکرد بهینه دست نمی یابد. روش کارآمد انتخاب نمونه های آموزشی نیازمنده بوجود آمدن است.

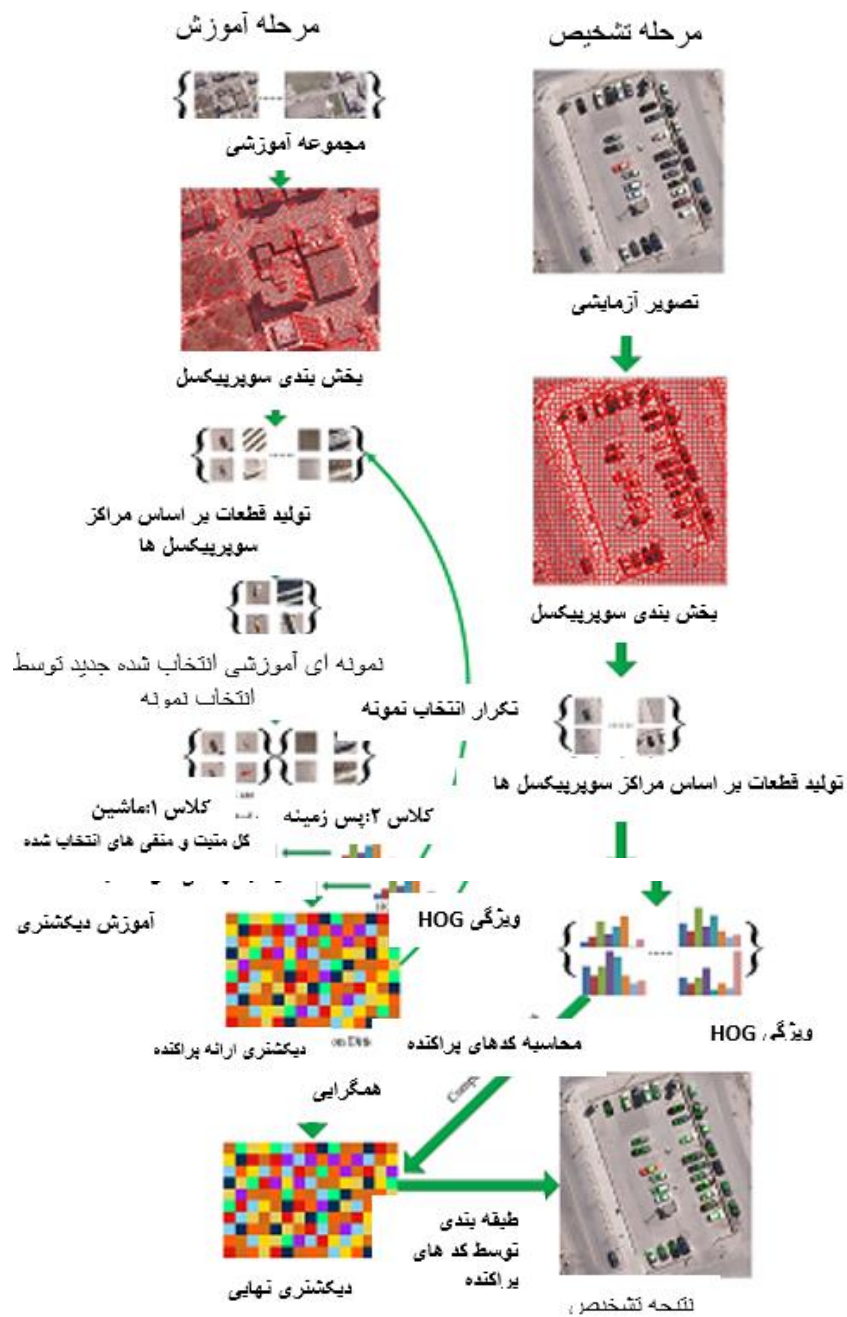
در نتیجه، نیاز زیادی برای استخراج روشی برای حل 2 مورد بالا وجود دارد.

3. راه حل پیشنهادی

A. چارچوب

همانگونه که در شکل 1 نشان داده شده است چارچوب کلی روش پیشنهادی ما شامل 2 مرحله آموزش فرهنگ جامع و تشخیص ماشین است. در مرحله آموزش، تصاویر آموزشی ابتدا توسط روش تولید سوپرپیکسل به سوپرپیکسل ها تقسیم می شود. بر اساس مراکز سوپرپیکسل ها، قطعات مهم را به عنوان مجموعه آموزشی تولید میکنیم. سپس زیرمجموعه آموزشی کوچکی برای شروع اولیه فرهنگ پراکنده کوچک انتخاب میکنیم. در روش ما، شبکه توصیف کنندگان HOG قطعات، به عنوان ورودی فرهنگ و دیکشنری استخراج می شوند. توسط دیکشنری آموزش یافته، شباهت بین نمونه های آموزشی باقی مانده و ماشین ها را تخمین میزنیم. نمونه های منفی با بالاترین شباهت و نمونه های مثبت با کمترین شباهت برای افزوده شدن به زیرمجموعه آموزشی انتخاب می شوند تا یک دیکشنری جدید را برای تکرار بعدی انتخاب نمونه ها آموزش دهد. تکرار انتخاب نمونه های آموزشی تا رسیدن به همگرایی ادامه می یابد. در این مقاله 2 وضعیت به عنوان حالات همگرایی مد نظر قرار گرفته اند. اول دیکشنری آموزش یافته شامل بیش از 2000 ایتیم و جز است. دوم دقت طبقه بندی تصاویر آزمایشی در نرخ فراخوانی 0.7 بیش از 80 درصد است. زمانی که همگرایی اتفاق افتاد، دیکشنری ارائه پراکنده را برای شناسایی وسایل نقلیه بکار میگیریم.

در مرحله تشخیص، تصویری آزمایشی ابتدا به سوپر پیکسل هایی که بر اساس مراکز آن ها تصویر آزمایشی را با کارایی بالا اسکن می کنیم، بخش بندی می شوند. بر اساس کد های پراکنده در طی اسکن کردن، کاندید های قطعات به ماشین ها و پس زمینه طبقه بندی می شود.



B. بخش بندی سوپریکسل

در روش پیشنهادی ما، بخش بندی سوپریکسل یک مرحله مهم است. شکست بخش بندی سوپریکسل تعریف شده به عنوان قطع ارتباط بین بخش بندی ها، دقت مکان اسکن کردن و عملکرد تشخیص را تحت تاثیر قرار می دهد. برای

بدست آوردن بخش بندی سوپرپیکسل با شکست کم ، روش بخش بندی سوپرپیکسل که مخصوص چارچوب خودمان طراحی شده را ارائه کردیم.

با داشتن جز واحد $P = \{p_i\}_{i=1}^j$ از تصویر $C = \{r(e), g(e), b(e)\}$ که e عضوی از C بوده و j شماره جز اولیه را نمایش میدهد ، $r(e), g(e), b(e)$ به ترتیب اجزای قرمز و سبز و آبی فضای رنگی پیکسل e را نمایش می دهند. سپس مرکز رنگی جز P_i توسط رابطه زیر محاسبه می شود:

$$R_{P_i} = \frac{1}{|p_i|} \sum_{e \in p_i} R(e)$$

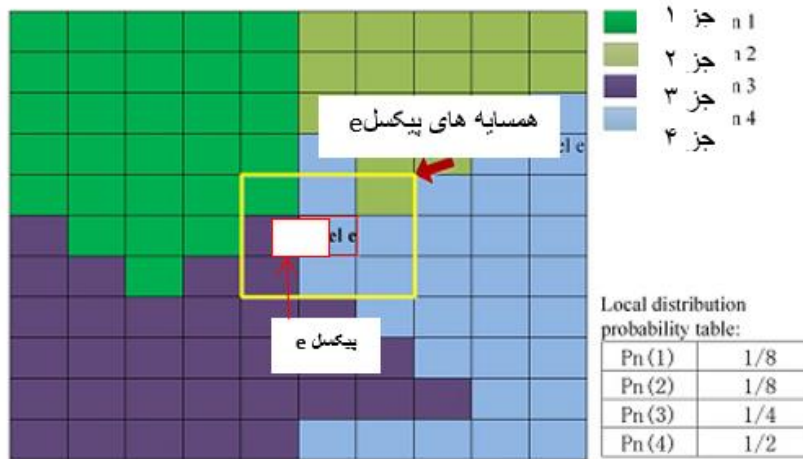
$$G_{P_i} = \frac{1}{|p_i|} \sum_{e \in p_i} G(e)$$

$$B_{P_i} = \frac{1}{|p_i|} \sum_{e \in p_i} B(e)$$

که در آن IP_i تعداد پیکسل های جز P_i است. $R_{p_i}, G_{p_i}, B_{p_i}$ مراکز رنگ اجزای RGB را به ترتیب نمایش می دهد. در مرحله بعدی ، بصورت تکراری محدوده های پیکسل هر جز را بر اساس 3 اندازه گیری بروز رسانی می کنیم. اولین اندازه گیری فاصله رنگ بین پیکسل مرزی e و مرکز رنگ جز همسایه P_i است که توسط رابطه زیر تعیین می شود:

$$d_c(e, p_i) = \frac{\sqrt{(r(e) - R_{p_i})^2 + (g(e) - G_{p_i})^2 + (b(e) - B_{p_i})^2}}{A_N} \quad (2)$$

که در آن A_N عبارت نرمالیزه کردن برای برابر 1 کردن کوچک ترین فاصله رنگی است . بنابراین A_N به عنوان کوچک ترین فاصله رنگی استخراج شده از بین فواصل بین پیکسل های مرزی و اجزای همسایگی متناسب تعیین می شود. بصورت تجربی ، قرار دادن مقدار 25 برای A_N برای بدست آوردن نتیجه خوب کافی است. دومین اندازه گیری فاصله فضایی بین پیکسل مرزی E و مرکز فضایی جز p_i در همسایگی اش است.



شکل ۲. امار توزیع احتمال در مکان محلی ۳*۳ پیکسل e

مرکز فضایی جز P_i بصورت زیر تعریف شده است:

$$S_{P_i}(u, v) = \left(\sum_{e \in P_i} w_e \cdot e_u, \sum_{e \in P_i} w_e \cdot e_v \right)$$

$$\hat{o}_e = \frac{R_{P_i}}{|r(e) - R_{P_i}|} + \frac{G_{P_i}}{|g(e) - G_{P_i}|} + \frac{B_{P_i}}{|b(e) - G_{P_i}|}$$

$$o_e = \frac{\hat{o}_e}{\sum_{e \in P_i} \hat{o}_e} \quad (3)$$

که در آن u و v مختصات مکان پیکسل e در تصویر و O_e و O_e وزن های پیکسل e و O_e عبارت نرمالیزه شده O_e است. سپس فاصله فضایی پیکسل e از جز همسایه P_i به طریق زیر محاسبه می شود:

$$d_s(e, p_i) = \frac{\sqrt{(u - u_{p_i})^2 + (v - v_{p_i})^2}}{B_N} \quad (4)$$

که B_n عبارت نرمالیزه سازی برای برابر 1 کردن کوچکترین فاصله فضایی است. بنابراین B_n به عنوان کوچکترین فاصله فضایی استخراج شده از فواصل بین پیکسل های مرزی و اجزای همسایه متناسب آنها تلقی می شود. بصورت تجربی مقدار دهی 10 به B_n برای بدست آوردن نتایج خوب کافی است.

سومین اندازه گیری، اطلاعات محلی پیکسل مرزی e است که احتمال آماری پیکسل هایی که به جز P_i در محدوده به مرکزیت e محاسبه می کند. مفهوم و رابطه به صورت زیر بیان می شود:

$$h(e, p_i) = \frac{N(e, p_i)}{L}$$

که در آن $N(x, P_i)$ تعداد پیکسل های در جز P_i در محدوده محلی e را نمایش می دهد و L تعداد کل پیکسل ها در محدوده محلی تعریف شده e است. یک مقدار بزرگ L استقامت اطلاعات محلی را بهبود می بخشد ولی پیچیدگی محاسبات را افزایش می دهد. در روش ما، یک محدوده محلی $3*3$ استفاده می کنیم. همان گونه که در شکل 2 نشان داده شده مربع زرد درون پنجره $3*3$ به عنوان محدوده محلی e تعریف شده است. احتمالات آماری پیکسل ها در جزیهای مجاور که جزیهای 1 تا 4 می شوند را محاسبه می کنیم. نهایتاً احتمال e که عضوی از P_i است بصورت زیر محاسبه می شود:

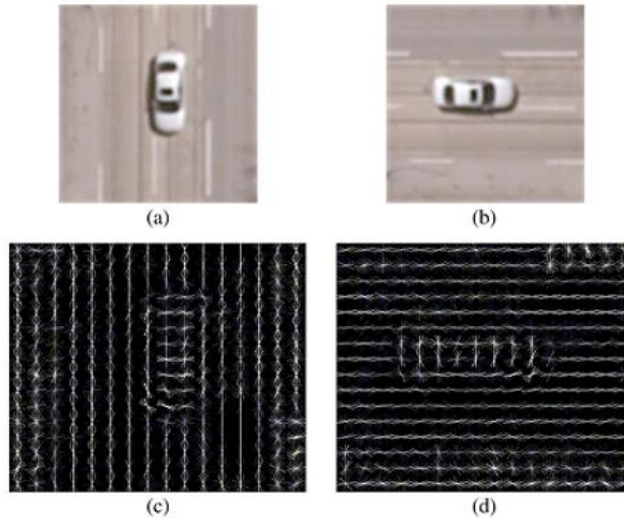
$$prob(e, p_i) = \frac{1}{d_c(e, p_i)} \cdot \frac{1}{d_s(e, p_i)} \cdot h(e, p_i).$$

بر اساس رابطه بالا، بصورت تکراری پیکسل های مرزی جز های بخش بندی شده را بروز رسانی میکنیم تا زمانی که پیکسل های مرزی بدون تغییر باقی بمانند یا به وضعیت پایان تکرار برسند. با استفاده از اطلاعات محلی، قطعات بخش بندی شده هموار بوده و در قطعی اتصالات به خوبی بکار گرفته شده اند. مزیت دیگر اطلاعات محلی این است که استقامت در برابر نویز بهبود یافته است. تحلیل های تفصیلی در بخش آزمایش مورد بحث قرار گرفته اند.

C. شبکه های توصیف گر های HOG

در این مقاله ما از شبکه های توصیف گر HOG برای شرح یک قطعه و تکه استفاده می کنیم. اینجا، یک مقدمه خلاصه در مورد شبکه های توصیف گر HOG پیشنهاد شده توسط Datal&Triggs ارائه می کنیم. با داشتن یک قطعه تصویر، ابتدا قطعه را به محدوده های فضایی کوچک تقسیم میکنیم (سلول ها). برای هر سلول، هیستوگرام محلی تک بعدی مسیر های گرادیان یا گرایش های لبه ای بر روی پیکسل های سلول فراهم می آوریم. سپس، مطالب هیستوگرام هر سلول برای تشکیل ارائه و تمثیل یک قطعه ترکیب می شوند. برای تغییر ناپذیری روشنایی، سایه بودن و غیره بهتر است وضوح پاسخ محلی را قبل از استفاده از آنها نرمالیزه کنیم. شکل 3 a و b قطعات تصویر یکسان با گرایش های متفاوت هستند. توصیف گر های HOG 2 قطعه با اندازه سلول 5 پیکسل را محاسبه کردیم. شکل 3 c و d به ترتیب توصیف گر های HOG شکل 3 a و b هستند. در شکل 3، توصیف گر های 2 قطعه متفاوت

هستند. همانگونه که در شکل 3 نشان داده شده، اگرچه به موفقیت بزرگی در زمینه تشخیص اشیا دست یافته است شبکه های HOG حساس به گرایش و سمت و سو هستند. توجه داشته باشید که بجای تصاویر RGB، ویژگی های مبتنی بر HOG را در تصاویر سیاه و سفید محاسبه کردیم.



با و b دو قطعه تصویر اصلی. c و d توصیف گر های HOG شکل a و b.

D. نمایندگی پراکنده

در روش ما، روش نمایش پراکنده ارائه شده توسط جیانگ و همکاران را اعمال می کنیم. [50] برای انتخاب نمونه، آموزش و آزمایش.

$Y = \{y_1, \dots, y_n\}$ که $R^{n \times N}$ است تعداد N سیگنال ورودی n بعدی را مشخص کند. سپس یادگیری یک دیکشنری قابل بازسازی با k عنصر برای ارائه پراکنده Y ما حل مسئله زیر بدست می آید:

$$\langle D, X \rangle = \arg \min_{D, X} \|Y - DX\|_2^2 \quad \text{s.t.} \quad \forall i, \|x_i\|_0 \leq T \quad (7)$$

که در آن T ترشلد پراکنده، D دیکشنری ارائه پراکنده، و X نمایان گر کد های پراکنده است. رابطه 7 با یک مسئله L1-عبارت می تواند جایگزین شود.

$$\langle D, X \rangle = \arg \min_{D, X} \|Y - DX\|_2^2 + \gamma \|X\|_1 \quad (8)$$

که Y پارامتری برای تعادل خطای بازسازی و کدهای ارائه پراکنده است. برابری روابط 7 و 8 در مرجع 51 ثابت شده بود. الگوریتم تجزیه k مقدار منحصر بفرد (K-SVD) رویکردی تکراری برای کاهش دادن تا حد امکان انرژی در رابطه 8 است و یک دیکشنری قابل بازسازی برای ارائه پراکنده می آموزد. بر عکس با داشتن دیکشنری D ، ارائه پراکنده X_i از سیگنال ورودی y_i به طریق زیر محاسبه می شود:

$$x_i = \arg \min_x \|y_i - Dx\|_2^2 + \gamma \|x\|_1. \quad (9)$$

با توجه به تفاوت کدهای پراکنده در بین دیگر کدها، کدهای پراکنده مستقیماً برای طبقه بندی قابل استفاده هستند. الگوریتم تعقیب متعامد متناسب برای حل رابطه 9 استفاده شده است. برای افزایش تفاوت و تمییز در کدهای پراکنده بدست آمده، عبارتی نمایانگر اطلاعات برچسب نمونه های آموزشی به دیکشنری آموزشی افزوده شده است. بنابراین تابع هدف برای تشکیل دیکشنری بصورت زیر تعریف می شود:

$$\langle D, A, X \rangle = \arg \min_{D, A, X} \|Y - DX\|_2^2 + \alpha \|Q - AX\|_2^2 \quad \text{s.t.} \quad \forall i, \|x\|_0 < T \quad (10)$$

که a توزیع نسبی بین بازسازی و تنظیم ثبات بر چسب را کنترل می کند، $Q = \{q_1, \dots, q_n\}$ عضو $R^{k \times n}$ کدهای پراکنده متمایز کننده سیگنال های ورودی Y برای طبقه بندی هستند، و A یک ماتریس تبدیل خطی است که کدهای پراکنده اصلی را به متمایزترین کدها در فضای ویژگی پراکنده R^k تبدیل می کند. عبارت $\|Q - AX\|_2^2$ خطای کدهای پراکنده متمایز کننده را نشان می دهد که سیگنال های همان کلاس و طبقه را مجبور میکند که ارائه پراکنده شبیه تری داشته باشند که عملکرد طبقه بندی بهتری را نتیجه می دهد. می گوئیم که

کدهای پراکنده متمایز کننده متناسب با سیگنال $q_i = [q_i^1, \dots, q_i^k]^t = [0, \dots, 1, 1, \dots, 0]^t \in R^k$

y_i است اگر مقادیر غیر صفر q_i در همان شاخص هایی اتفاق بیافتند که سیگنال ورودی y_i و عبارت دیکشنری dk برچسب یکسانی را داشته باشند. برای مثال فرض کنید $D = \{d_1, \dots, d_6\}$ و $Y = \{y_1, \dots, y_6\}$ که y_1, y_2, d_1, d_2 از کلاس 1 و y_3, y_4, d_3, d_4 از کلاس 2 و y_5, y_6, d_5, d_6 از کلاس 3 هستند. سپس Q بصورت زیر تعریف

می شود:

$$Q \equiv \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

که هر ستون متناسب با کدهای پراکنده متمایز کننده برای هر سیگنال ورودی است.

در طی یادگیری دیکشنری، عبارتی که قدرت متمایز کنندگی کدهای پراکنده را بین طبقات مختلف بهینه می سازد، در تابع هدف یادگیری دیکشنری گنجانده شده است. سپس تابع هدف بصورت زیر نشان داده شده است:

$$\langle D, W, A, X \rangle = \arg \min_{D, W, A, X} \|Y - DX\|_2^2 + \alpha \|Q - AX\|_2^2 + \beta \|H - WX\|_2^2 \quad \text{s.t.} \quad \forall i, \|x_i\|_0 < T \quad (11)$$

که عبارت $\|H - WX\|_2^2$ خطای طبقه بندی را نشان می دهد. W پارامترها و عوامل طبقه بندی کننده را نشان می دهد. $H = \{h_1, \dots, h_N\}$ که عضوی از $\mathbb{R}^{m \times N}$ است، برچسب های طبقه ها و کلاس های سیگنال های ورودی Y است. $h_i = \{0, 0, \dots, 1, \dots, 0, 0\}^T$ که عضوی از \mathbb{R}^m است بردار برچسب متناسب با سیگنال ورودی y_i است. m تعداد طبقات و a و B ضرایبی هستند که توزیع نسبی عبارات متناسب را کنترل می کنند. برای بکار گیری الگوریتم K-SVD برای حل 11، رابطه 1 به صورت زیر دوباره نویسی می شود:

$$\langle D, W, A, X \rangle = \arg \min_{D, W, A, X} \left\| \begin{pmatrix} Y \\ \sqrt{\alpha}Q \\ \sqrt{\beta}H \end{pmatrix} - \begin{pmatrix} D \\ \sqrt{\alpha}A \\ \sqrt{\beta}W \end{pmatrix} X \right\|_2^2 \quad \text{s.t.} \quad \forall i, \|x_i\|_0 \leq T. \quad (12)$$

توجه کنید که $Y_{\text{new}} = (Y^t, \sqrt{\alpha}Q^t, \sqrt{\beta}H^t)^t$ بوده و $D_{\text{new}} = (D^t, \sqrt{\alpha}A^t, \sqrt{\beta}W^t)^t$ است. سپس رابطه 12 بصورت تابع زیر در می آید:

$$\langle D_{\text{new}}, X \rangle = \arg \min_{D_{\text{new}}, X} \{\|Y_{\text{new}} - D_{\text{new}}X\|_2^2\} \quad \text{s.t.} \quad \forall i, \|x_i\|_0 \leq T. \quad (13)$$

رابطه 13 به صورتی که الگوریتم K-SVD حل می کند است. بعد از بدست آوردن مجموعه های D, A, W از مجموعه D_{new} نمی توانیم فقط D, A, W را برای آزمایش استفاده کنیم چرا که آنها بطور مشترک L2-نرمالیزه شده اند که بصورت $\forall k, \|d_k^t, \sqrt{\alpha} a_k^t, \sqrt{\beta} w_k^t\|_2 = 1$ است. بنا بر این دیکشنری مطلوب D ، پارامتر های تبدیل A و

پارامتر های طبقه بندی کننده W بصورت زیر پیشنهاد می شوند:

$$\hat{D} = \left\{ \frac{d_1}{\|d_1\|_2}, \dots, \frac{d_K}{\|d_K\|_2} \right\}, \hat{A} = \left\{ \frac{a_1}{\|a_1\|_2}, \dots, \frac{a_K}{\|a_K\|_2} \right\}$$

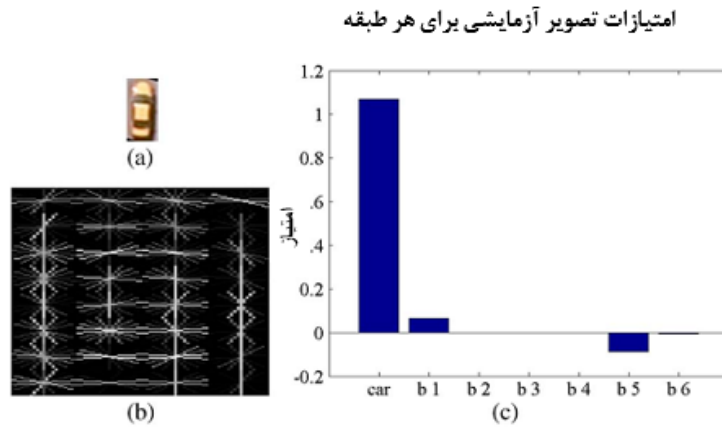
$$\hat{W} = \left\{ \frac{w_1}{\|w_1\|_2}, \dots, \frac{w_K}{\|w_K\|_2} \right\}. \quad (14)$$

D, A, W مطلوب به طور مستقیم به آزمایشات اعمال می شوند. پیش بینی طبقه بندی نهایی l می تواند به صورت زیر نمایش داده شود:

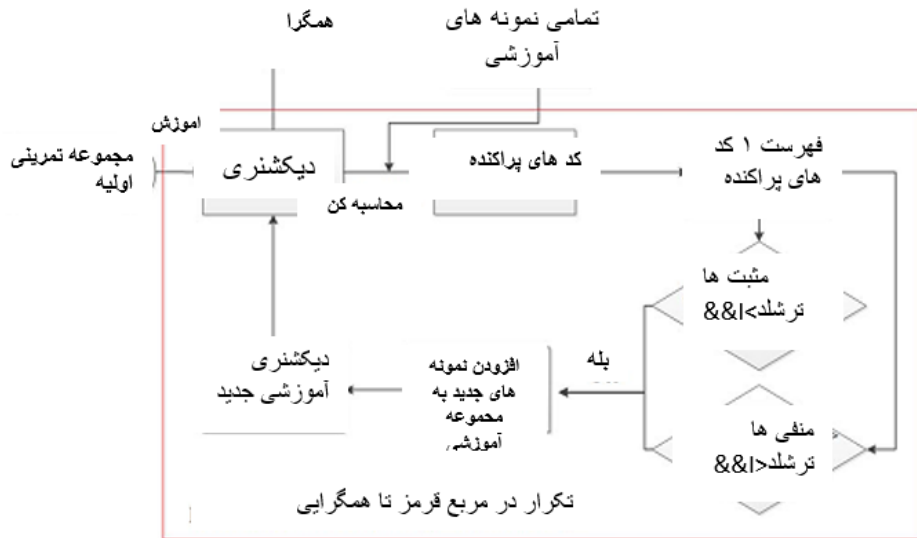
$$l = \hat{W} \hat{x}_i.$$

برچسب y_i به عنوان امتیازات طبقه بندی متناسب با هر کلاس و طبقه فرض شده است. در روش ما، شبکه ای ویژگی HOG به عنوان سیگنال های ورودی اولیه Y بکار گرفته شده اند. بصورت نشان داده شده در شکل 4، شبکه ویژگی های HOG یک نمونه ماشین را به عنوان آزمایش ورودی استخراج می کنیم و کد های پراکنده نمونه آزمایشی را حل می کنیم. بنا به کدهای پراکنده، امتیازات قطعه آزمایشی را متناسب با هر طبقه محاسبه می کنیم. واضحاً، نمونه آزمایشی دارای بالاترین امتیاز برای ماشین ها است. بنا بر این نمونه آزمایشی را به عنوان یک ماشین طبقه بندی می کنیم. طبقات پس زمینه b1-b6 توسط پروسه انتخاب نمونه ما تشکیل یافته و بر اساس شباهت های تخمین زده شده به وسایل نقلیه بر خلاف طبقات پس زمینه واقعی آنها مانند علف ها، زمین، و تهویه گر هوا ساخته شده اند.

شکل 4. (a) نمونه آزمایشی ماشین ها. (b) ویژگی HOG نمونه آزمایشی. (c) امتیازات طبقه نمونه آزمایشی در پاسخ به هر طبقه بر اساس کد های پراکنده آنها. b1-b6 طبقات پس زمینه هستند.



شکل 4. (a) نمونه آزمایشی ماشین ها. (b) ویژگی HOG نمونه آزمایشی. (c) امتیازات طبقه نمونه آزمایشی در پاسخ به هر طبقه بر اساس کد های پراکنده آنها. b1-b6 طبقات پس زمینه هستند.



شکل 5. چارچوب روش پیشنهادی انتخاب نمونه آموزشی

E. انتخاب نمونه های آموزشی تکراری

در این بخش رویکرد انتخاب خودکار نمونه های آموزشی را معرفی می کنیم که تفاوت درون و کامل بودن درون را در نظر می گیرد تا مجموعه آموزشی فشرده برای بهبود کارایی طبقه بندی و آموزش تشکیل دهد.

شکل 5 چارچوب روش انتخاب نمونه های آموزشی تکراری را نشان می دهد. ابتدا بصورت دستی ده ها نمونه از ماشین ها و نمونه های پس زمینه برای مقدار دهی اولیه به مجموعه آموزشی انتخاب می کنیم. دوم شبکه های ویژگی HOG از نمونه های آموزشی به عنوان سیگنال ورودی Y برای آموزش دیکشنری بنا بر رابطه 11 استخراج می شوند. سوم ، کد های پراکنده تمامی نمونه های آموزشی محاسبه می شوند. هر نمونه به فهرست توزیع طبقه ای تخصص می یابد. بر اساس این فهرست ، نمونه ها را برای افزودن به مجموعه آموزشی انتخاب می کنیم. فهرست توزیع طبقه I یک نمونه بصورت زیر تعریف می شود:

$$I = \frac{l_i}{\sum_{i=1}^m |l_i|}$$

که 1 اطلاعات برجسته نمونه ها که از 15 محاسبه شده است و m تعداد کل طبقات است. بر اساس فهرست توزیع طبقات ، شباهت بین یک نمونه و یک ماشین را تخمین می زنیم. چهارم ، نمونه ها بری پیوستن به مجموعه آموزشی بر اساس 2 ضابطه زیر انتخاب می شوند:

1. یک نمونه مثبت که مقدار I آن کمتر از ترشلد تعریف شده است، به عنوان یک نمونه مثبت جدید انتخاب می شود. ترشلد از ترشل کوچک (0.1 در روش ما) شروع می شود. سپس پیوسته در طی تکرار انتخاب نمونه ها افزایش می یابد.

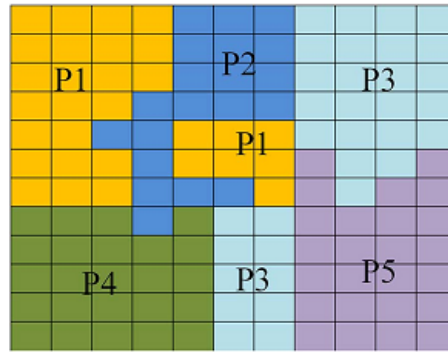
2. یک نمونه منفی که مقدار I از ترشلد معین بیشتر است به عنوان نمونه منفی جدید انتخاب می شود. ترشلد از ترشلد بزرگ (0.9 در روش ما) شروع می شود. سپس در طی تکرار انتخاب نمونه ها کاهش می یابد.

ضابطه اول تفاوت بین نمونه های درونی ماشین را تضمین می کند و دومین ضابطه قابلیت تمایز نمونه های آموزشی انتخاب شده برای طبقه بندی ماشین ها و پس زمینه را تضمین می کند. برای جلوگیری از انتخاب دوباره نمونه ها، تمامی نمونه های انتخاب نشده را صفر و نمونه های انتخاب شده را 1 برجسته می کنیم. پنجم، دیکشنری جدید و بزرگ تر بر اساس مجموعه آموزشی جدید آموزش داده شده است. ما بصورت تکراری این مراحل را برای انتخاب نمونه ها اجرا می کنیم تا زمانی که شرایط همگرایی ذکر شده بدست می آیند.

همان طور که ذکر شد ، شبکه های توصیف گر HOG استفاده شده در روش ما حساس به گرایش هستند. گرایش وسیله نقلیه در تصویر آزمایشی نا مشخص است. بنابراین تصویر آزمایشی باید در جهات مختلف اسکن شود. معمولا هر قطعه آزمایشی با حداقل فاصل زاویه 5 درجه (یا 30 درجه یا 45 درجه و غیره) برای اسکن کردن چرخش پیدا می کند که منجر به کاهش چشمگیر کارایی تشخیص می شود. برای بهبود بیشتر کارایی تشخیص ، گرایش و جهت اصلی هر قطعه را حدس می زنیم و بصورت خودکار تمامی قطعات را به جهت اصلی خودشان چرخش می دهیم تا پایداری جهت را در مدت تشخیص وسیله نقلیه و آموزش دیکشنری ایجاد کند. توسط تخمین جهت اصلی ، تنها به بررسی هر قطعه آزمایشی در یک جهت نیاز داریم. بنابراین ما کارایی تشخیص را بهبود دادیم. با در نظر گرفتن اینکه یک قطعه ماشین معمولا دارای 2 خط موازی طولانی در امتداد جهت طولی ماشین است ، جهت اصلی قطعه را جهت خطوط موازی با طول درازتر از ترشلد معین تعیین می کنیم. در این مقاله ، تبدیلات Canny و Hough برای تشخیص لبه و تشخیص خط استفاده شده اند. با تاثیر از تعیین روشنایی ، انسداد ها و نویز ها ، خطوط موازی استخراج شده شامل خطوطی می شوند که مربوط به جهت طولی ماشین نمی شوند. برای کاهش هم پوشانی خطوطی که جهات متفاوتی با طول ماشین دارند ، ما خطوط موازی طولانی تشخیص داده شده را به طبقات مختلف بر اساس زوایای آنها خوشه بندی می کنیم. زاویه میانگین خوشه ای که بیشتری خطوط را داراست به عنوان جهت اصلی قطعه در نظر گرفته می شود. شکل 6 نتایج قطعات بعد از چرخش بر اساس جهت اصلی آنها را نشان می دهد. همان طور که نشان داده شده ، روش ما بصورت موثر مشکل چرخش قطعات را حل می کند و بسیاری از ماشین ها را در جهت عمودی خودشان بصورت صحیح چرخش می دهد.



شکل 6. نمونه ها بعد از چرخش به جهت اصلی خودشان



شکل ۷. نمونه ای از شکست بخش بندی. P1-5 نمایان گر جز های تصویر هستند. در شکل P1 و P3 بخش دارند که جدا از هم هستند. وضعیت P1 و P3 شکست بخش بندی نامیده می شوند.

4. آزمایشات و بحث ها

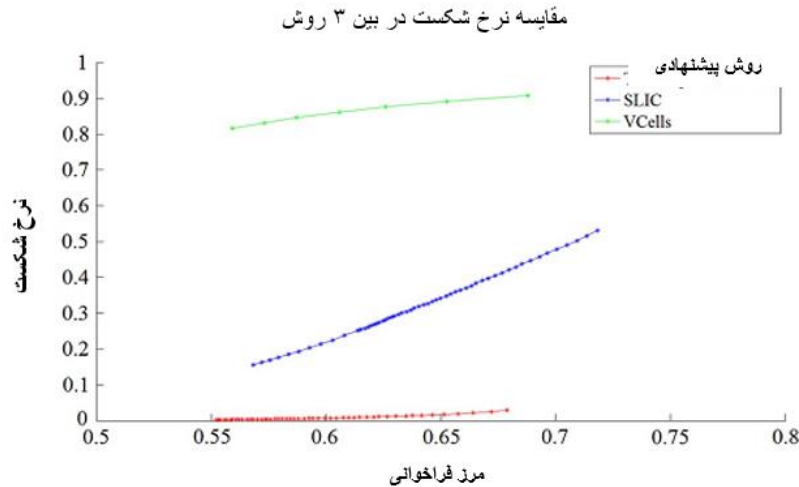
در این قسمت ابتدا بحثی در مورد بخش بندی سوپرپیکسل فراهم می کنیم. سپس الگوریتم خود را در 2 مجموعه داده که مجموعه داده Toronto و OIRDS هستند، آزمایش می کنیم. در هر دو مجموعه داده، حقایق زمین تصاویر آزمایشی توسط محدوده های مربعی در اطراف ماشین ها مشخص شده و برچسب زده شده بودند. تنها تشخیص هایی که دقیقاً در حقیقت زمین قرار دارند، تشخیص های صحیح و درست در نظر گرفته می شوند. اگر یک حقیقت زمین دفعات زیادی دوباره تشخیص داده شود، فقط یکی از آنها تشخیص مثبت صحیح در نظر گرفته می شود. بنابراین تشخیص های همپوشانی به عنوان هشدار های اشتباه در نظر گرفته می شوند.

A. بحث بخش بندی سوپرپیکسل

در این بحث، ما روش بخش بندی سوپرپیکسل خود را در پایگاه داده عمومی Berkeley آزمایش کردیم که این پایگاه داده شامل 300 عکس می شود. برای تحلیل استقامت روش خودمان در برابر شکست، که به عنوان قطع ارتباط بخش ها تعریف می شود، نرخ شکست روش خود بر روی پایگاه داده Berkeley را ارزیابی کردیم. شکل 7 مثالی از شکست بخش بندی را نشان می دهد. در شکل یک تصویر به 5 قسمت به نام های P1-5 بخش بندی شده است. ولی قسمت های P1 و P3 دارای 2 قسمت جدا از هم هستند که به عنوان شکست بخش بندی تلقی می شود. با داشتن بخش بندی P، نرخ شکست برابر است با:

$$BR = \frac{\varphi(P)}{|P|}$$

که $\varphi(P)$ نمایان گر تعداد بخش های جدای P است و IPI نمایان گر تعداد کل بخش های P است.



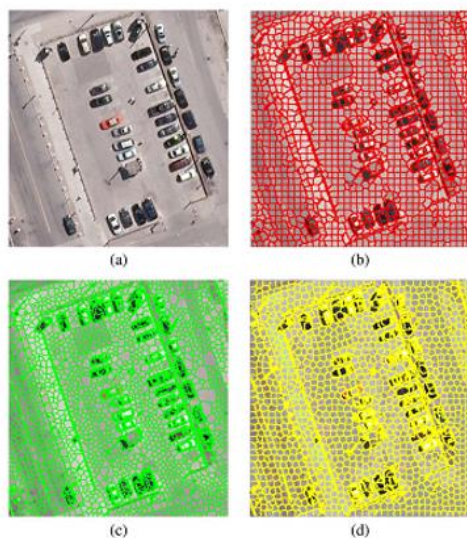
شکل ۸. مقایسه نرخ شکست در نتایج بخش بندی توسط روش ما، VCells، و SLIC در پایگاه داده Berkeley. خطوط قرمز، آبی و سبز نشانگر به ترتیب روش ما، SLIC، و VCells هستند.

برای مقایسه، ما همچنین نرخ شکست Vcells و SLIC بر روی این پایگاه داده را ارزیابی کردیم. کد های Vcell ها از نویسنده بدست آمده بودند و کد های SLIC از VLFeat بدست آمده است. در آزمایش ما، ما اندازه سوپرپیکسل را در حدود 200 پیکسل بر سوپرپیکسل ثابت کردیم و نرخ فراخوانی در محدوده 55درصد تا 70 درصد نگه داشتیم. فقط پیکسل مرزی بخش بندی شده که در مرز های حقیقت زمین قرار دارد، به عنوان پیکسل مرزی بخش بندی شده صحیح در نظر گرفته شده بود. شکل 8 نتایج آزمایش را نشان می دهد؛ محورهای افقی و عمودی نرخ فراخوانی و نرخ شکست را به ترتیب نشان می دهند. در شکل، خط قرمز نشانگر روش ما است که نشان می دهد روش ما کمترین نرخ شکست را بین 3 روش دارد. هر دوی VCells و SLIC نرخ های شکست بالایی در حالی که مرز فراخوانی بیشتر از 0.55 بود، بدست آوردند. شرایط برای VCells بدتر بود. بر عکس در مدت بخش بندی، ما از اطلاعات محلی برای بهبود پروسه بخش بندی استفاده کردیم و بصورت موفقیت آمیزی بخش بندی را با نرخ شکست پایین کنترل کردیم. شکل 9 مقایسه بصری بخش بندی در تصویر هوایی توسط روش ما، VCells و SLIC را نشان می دهد. پارامترهای بخش بندی همه روش ها در آزمایشات در پایگاه داده Berkeley یکسان نگه داشته شده

بودند.مشاهده می شود که نتایج بخش بندی روش ما صاف تر و مظم تر دو روش دیگر است.زمانی که اهدافی با محتویات پیچیده مانند ماشین ها بکار گرفته می شوند، نتایج بخش بندی VCells و SLIC شکست آنها را نشان داد که منجر به تولید قطعات بخش بندی بسیار کوچک شد.این موضوع نه تنها بار وظیفه تشخیص بلکه نرخ تشخیص اشتباه را نیز افزایش می دهد.

B. مجموعه داده تورنتو

ما عملکرد الگوریتم خود را در تصویر هوایی که شهر تورنتو را پوشش می دهد با سایزو اندازه 11500 پیکسل در 7500 پیکسل و عمق رنگ 24 بیت /پیکسل (RGB) آزمایش کردیم.وضوح فضایی تصویر هوایی 0.15 متر است که در این وضوح ، یک ماشین شامل 38*16 پیکسل می شود.در آزمایش ما ، تصویر را به زیرمحدوده ها و زیرمحدوده های انتخاب شده برای آموزش و آزمایش برش می دهیم.شکل 10 یک زیرتصویر که یک پارکینگ را پوشش می دهد نشان می دهد.در آزمایش، 13 زیر تصویر برای آموزش و 8 تصویر برای آزمایش انتخاب شده اند.تعداد کل ماشین ها در این مجموعه 1589 است.عموما ، اندازه قطعه اسکن کننده بزرگتر از اندازه ماشین ها در تصاویر آزمایشی قرار داده می شود.در آزمایشات ما، اندازه قطعه اسکن کننده 41*21 پیکسل است.دلیل استفاده از این اندازه قطعه اسکن کننده این است که ما فقط ویژگی های محتویات و بافت وسایل نقلیه را در روش خودما در نظر میگیریم. اطلاعات متن پس زمینه نزدیک در کارهای آتی ما مطالعه خواهد شد.



شکل ۹. مقایسه نتایج بخش بندی سوپریکسل در بین روش پیشنهادی ما و VCells و SLIC. (a) تصویر هوایی اصلی با وضوح بالا. (b)-(d) نتایج بخش بندی روش پیشنهادی ما، VCells و SLIC به ترتیب. پارامترهای پیکربندی همه روش‌ها در پایگاه داده Berkeley یکسان بودند.



شکل 10. مثال پارکینگ از مجموعه داده تورنتو

Area	1	2	3	4	5	6	7	8	9	10	11	12	13
Width	1193	1758	1961	1629	1087	1374	1887	2086	1938	1523	2744	4060	1185
Length	775	1002	1136	1085	926	920	633	986	743	769	1264	2948	888

جدول 1 اندازه زیرمحدوده‌ها برای آموزش در مجموعه داده تورنتو

C. تشکیل مجموعه آموزشی کامل و فشرده

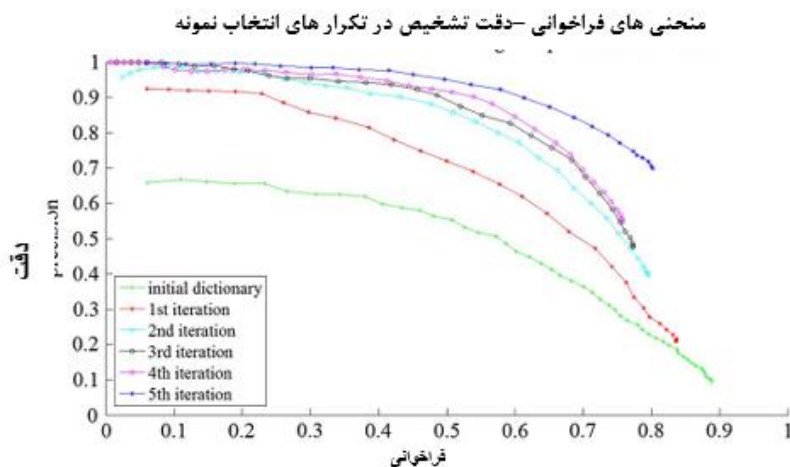
در آزمایشات ما، قطعات مثبت (ماشین‌ها) و قطعات منفی (پس زمینه) از 13 زیرمحدوده انتخاب کردیم تا دیکشنری ارائه‌پراکنده را آموزش دهیم. جدول 1 اندازه هر زیرمحدوده برای آموزش را نشان می‌دهد. ما تمامی زیر

محدوده ها رابه سوپرپیکسل هایی با اندازه حدودی 400 پیکسل بخش بندی کردیم که 184710 سوپر پیکسل را تولید کرد.بر همین اسا،184710 قطعه آموزشی که شامل 5169 قطعه ماشین بودند بر اساس مراکز سوپرپیکسل ها بدست آمده اند.انتخاب یک زیرمجموعه آموزشی کامل و کوچک ضروری است چرا که کل مجموعه آموزشی ایجاد شده برای آموزش بسیار بزرگ بوده و مقدار زیادی از اطلاعات اضافی و زائد هستند.ولی انتخاب دستی زیرمجموعه آموزشی کامل و کوچک مشکل است.برای ایجاد خودکار زیرمجموعه آموزشی کامل، ما روش انتخاب نمونه آموزشی خود را برای انتخاب نمونه های آموزشی نماینده وار بکار گرفتیم.به منظور کاهش پیچیدگی محاسباتی ، تمامی این قطعات ابتدا در جهت اصلی خود چرخش داده شدند. سپس هر قطعه به اندازه واحد $41*21$ پیکسل همانگونه که در شکل 6 نشان داده شده ، کاهش یافته است.

از قطعات آموزشی چرخش یافته و کاهش یافته ، بصورت دستی 60 قطعه ماشین و 120 قطعه پس زمینه را انتخاب می کنیم .برای مثبت ها ، ما قطعات ماشین با محتویات و بافت واضح بدون تداخل(مانند سایه ،انسداد و غیره) را انتخاب کردیم.برای منفی ها ، نمونه هایی که شبیه ماشین ها بودند را انتخاب کردیم.شبکه ویژگی های HOG به عنوان ورودی اولیه دیکشنری ارائه پراکنده استخراج شده بود.تمامی ویژگی های مثبت برچسب 1 و تمام ویژگی های منفی 1- برچسب زده شده اند.با استفاده از دیکشنری آموزش داده شده، کد های پراکنده و امتیازات طبقه بندی قطعات آموزشی باقی مانده را محاسبه کردیم.بر اساس امتیازات محاسبه شده ، بعضی قطعات برای ملحق شدن به زیرمجموعه آموزشی انتخاب شده قبلی انتخاب شده بودند.با زیرمجموعه آموزشی جدید ،دیکشنری جدید برای انتخاب نمونه در تکرار بعدی آموزش داده شده بود.ما انتخاب نمونه آموزشی را بعد از 5 تکرار متوقف کردیم بدلیل اینکه در آن زمان مقدار دقت طبقه بندی بیشتر از 80درصد در نر فراخوانی 0.7 را بدست آوردیم(که مقدار دقت تشخیص رضایت بخش در این مقاله است).نهایتا زیرمجموعه آموزش دیکشنری کامل و کوچک با 180 قطعه ماشین و 1080 قطعه پس زمینه ایجاد شد.

به دلیل تغییرات روشنایی و انسداد، جهت اصلی تخمین زده شده ممکن است جهت عمودی مطلوب ماشین نباشد که منجر به حذف تشخیص ماشین ها شود.بنابراین ، ما دو جهت از هر قطعه را اسکن کردیم.بعد از گروه بندی خطوط

موازی تشخیص داده شده، 2 گروهی که بیشترین تعداد خطوط موازی را نسبت به بقیه داشتند، انتخاب کردیم. جهت 2 گروه انتخاب شده به عنوان جهت اصلی تعیین شدند. یک قطعه آزمایشی در 2 جهت اصلی مورد آزمایش قرار گرفت. برای جل. گیری از تشخیص های دوباره، فقط نتیجه ای با امتیاز مثبت بالاتر به عنوان نتیجه آزمایش مدنظر قرار گرفت. یک قطعه آزمایشی بدون خطوط بدون هیچ جهت و گرایش در آزمایش ما مورد بررسی قرار گرفته بود.



شکل 11. نتایج تشخیص با دیکشنری های آموزش داده شده در هر تکرار. سبز، قرمز، فیروزه ای، سیاه و سرخابی

به ترتیب نتایج تکرار ها از الیه تا پنجم را نشان می دهد.

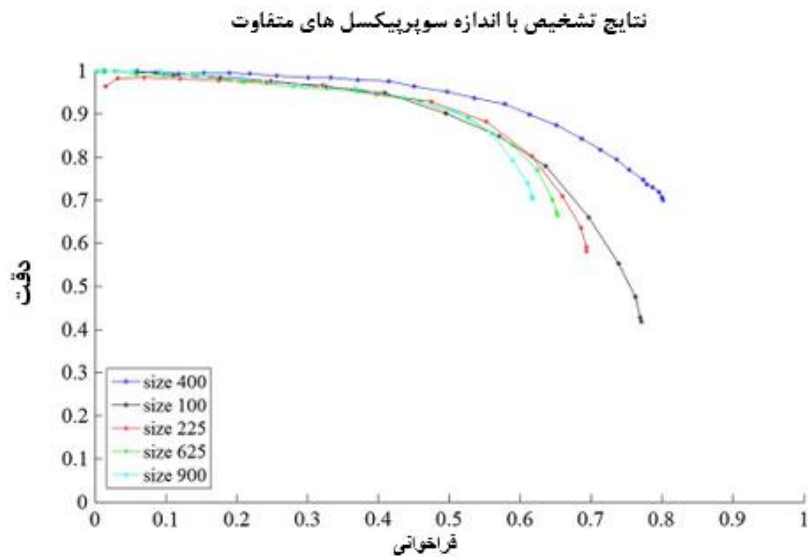
برای نشان دادن کارایی انتخاب نمونه آموزشی تکراری خودمان، از دیکشنری های آموزش یافته در طی تکرار ها برای تشخیص وسیله نقلیه در تصاویر آزمایشی استفاده کردیم. شکل 11 منحنی های دقت-فراخوانی تشخیص با استفاده از دیکشنری های آموزش یافته با نمونه های انتخاب شده در تکرار های مختلف یک تا پنج را نشان می دهد. دقت تشخیص به صورت یکنواخت در تکرار انتخاب نمونه ها افزایش می یابد بدلیل اینکه منفی هایی که به مثبت ها شباهت دارند یکی یکی به مجموعه آموزشی افزوده شده اند. با افزوده شدن منفی هایی که به مثبت ها شباهت دارند، مرز تشخیص بین منفی ها و مثبت ها بیشتر و بیشتر دقیق تر می شود. بنابراین توانایی جداسازی وسایل نقلیه، همانند منفی های آزمایشی که به وسایل نقلیه در ویژگی های فضایی شباهت دارند، در طی تکرار ها افزایش می یابد. سپس دقت تشخیص در طی تکرار ها افزایش می یابد. نتایج تجربی و آزمایش، کارایی روش ما برای تشکیل

خودکار مجموعه آموزشی کامل و کوچک را ثابت می کنند. در آزمایش ما، نتیجه رضایت بخشی بعد از 5 تکرار بدست آوردیم. در مقایسه بعدی نتیجه تکرار پنجم به عنوان نتیجه نهایی فرض شده است.

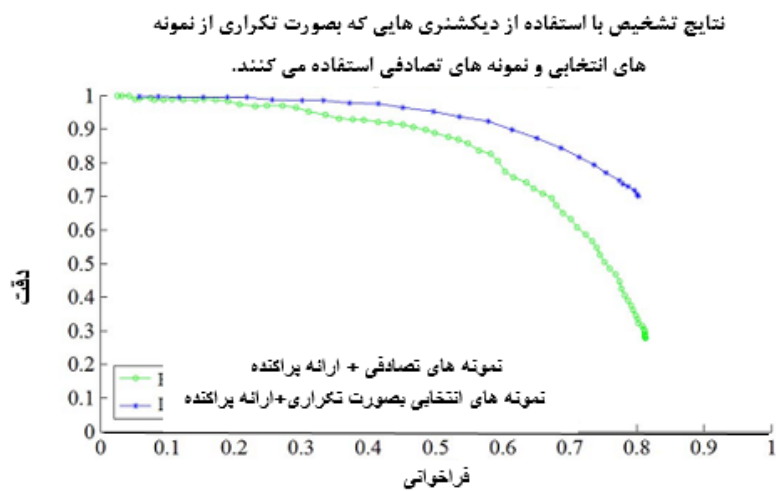
D. حساسیت اندازه سوپرپیکسل

بخش بندی سوپرپیکسل پروسه ای حیاتی برای تشخیص ماشین ها در تصاویر هوایی با وضوح بالا در روش ما است. اندازه سوپرپیکسل، دقت تشخیص و نرخ فراخوانی را تحت تاثیر قرار می دهد. اندازه بزرگ سوپرپیکسل مکان های تشخیص و هشدار های اشتباه را کاهش می دهد ولی تشخیص های حذفی را افزایش می دهد. در سوی دیگر، اندازه کوچک سوپرپیکسل بار تشخیص و دوباره تشخیص را افزایش می دهد ولی نرخ فراخوانی را بهبود می بخشد. این موضوع تعادلی در اندازه سوپرپیکسل برای تشخیص است.

در این مقاله، برای تعیین بهترین اندازه سوپرپیکسل، 5 اندازه سوپرپیکسل (100 و 225 و 400 و 625 و 900) را برای بخش بندی تصاویر برای آموزش و تشخیص آزمایش کردیم. از یک بخش بندی با اندازه سوپرپیکسل داده شده، نمونه های آموزشی را انتخاب و تصاویر را تشخیص دادیم. شکل 12 نتایج تشخیص با اندازه سوپرپیکسل های متفاوت در چارچوب ما را نشان می دهد. در شکل 12 نتیجه تشخیص با اندازه سوپرپیکسل 400 نسبت به بقیه برتری دارد. دقت تشخیص با اندازه های سوپرپیکسل 100 و 225 به سرعت کاهش می یابد بدلیل اینکه ماشین ها در تصاویر آزمایشی بیش از حد بخش بندی شده بودند که باعث افزایش دوباره تشخیص در طی تشخیص شد. در سمت دیگر، حذفیات زمانی که اندازه سوپرپیکسل 625 یا 900 بود افزایش یافت که منجر به نرخ فراخوانی ماکسیمم کوچک تر شد. برای تعادل تشخیص های دوباره و حذفیات تشخیص در مطالعه ما سایز سوپرپیکسل 400 در نظر گرفته شده است.



شکل 12. نتایج تشخیص در چارچوب پیشنهادی با اندازه های بخش بندی سوپرپیکسل متفاوت. اندازه سوپرپیکسل ها شامل 100 و 225 و 400 و 625 و 900 می شود.



شکل 13. مقایسه نتایج تشخیص با استفاده از دیشکنری های آموزش یافته تکراری و آموزش یافته تصادفی. خطوط آبی و سبز به ترتیب نشانگر نتایج نمونه های انتخابی بصورت تکراری و نتایج نمونه های انتخابی بصورت تصادفی هستند.

E. تاثیرات بر روی نمونه های آموزشی

برای تحلیل اثر نمونه های آموزشی بر روی دقت تشخیص وسایل نقلیه، ماشین ها را با دیکشنری آموزش یافته بصورت تکراری خود و دیکشنری آموزش یافته بصورت تصادفی تشخیص دادیم. هر یک از دیکشنری ها شامل 180 ماشین و 1080 اشیا سر از ماشین بودند. شکل 13 نتایج مقایسه را نشان می دهد. خطوط آبی و سبز منحنی های دقت-فراخوانی نتایج تشخیص با دیکشنری آموزش یافته بصورت تکراری خود و دیکشنری آموزش یافته بصورت تصادفی را نشان می دهند. هر دوی تشخیص ها در چارچوب پیشنهادهای تشخیص وسیله نقلیه بودند. همطور که در شکل 13 نشان داده شده است، روش دیکشنری آموزش یافته بصورت تکراری برای تشخیص ماشین را از دیکشنری آموزش یافته بصورت تصادفی بهتر عمل می کند. روش دیکشنری آموزش یافته بصورت تکراری دقت بالاتری برای هر مقدار نرخ فراخوانی دارد که کارایی استراتژی انتخاب نمونه آموزشی بصورت تکراری را ثابت می کند. مجموعه آموزشی انتخابی بصورت تصادفی توانایی دارا شدن تمامی نمونه های منفی نماینده وار را در مجموعه آموزشی اصلی برای آموزش دیکشنری با تمایز بالا را ندارد.

F. عملکرد در مجموعه داده تورنتو

مجموعه داده تورنتو برای اثبات و تصدیق عملکرد الگوریتم ما می شود. جدول 2 اندازه تصاویر آزمایشی را نشان می دهد. در این مقاله، تصاویر آزمایشی دقت فضایی یکسانی با تصاویر آموزشی دارند. 8 تصویر آزمایشی را به سوپرپیکسل هایی با اندازه 400 بخش بندی کردیم و نماینده های قطعات با اندازه $61*61$ پیکسل بر اساس مراکز سوپرپیکسل ها تولید شد. این قطعات در جهت اصلی خودشان چرخش یافته و به قطعاتی با سایز $41*21$ پیکسل بریده شده بودند. ما تنها بافت و محتویات موسایل نقلیه را در این مقاله برای روش پیشنهادی خود در نظر می گیریم. استفاده از اطلاعات پس زمینه در کار و مطالعه بعدی ما انجام خواهد گرفت. در حین تشخیص، کد های پراکنده نمایندگان قطعات برای طبقه بندی آنها در طبقات ماشین یا پس زمینه محاسبه شده بود. دیکشنری آزمایش ما، بعد از 5 تکرار انتخاب نمونه های آزمایشی بدست آمده بود. شکل 18 دو نتیجه تشخیص روش ما را نشان می دهد. در شکل 18،

خط قرمز تشخیص های اشتباه و خط سبز تشخیص های صحیح را نشان می دهند. نتایج ما عملکرد خوبی را نشان می دهند که دارای دقت و فراخوانی بالایی در مناطق شهری پیچیده است.

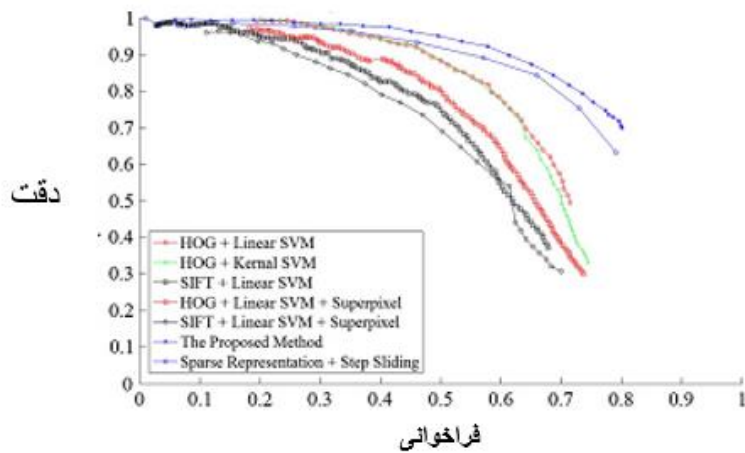
همچنین روش های محبوب دیگر را نیز بر روی تصاویر آزمایشی برای مقایسه انجام دادیم که شامل HOG+خطی SVM ، SVM + HOG هسته و مرکزی ، SVM + SIFT خطی می شوند. تمامی کد ها در آزمایش ما از منابع در دسترس عموم VLFeat بدست آمده اند. در این روش ها ، استراتژی اسکن پنجره لغزنده با گام لغزش 5 پیکسل در 2 محور عمودی و افقی ، استفاده شده بود. برای هر مکان اسکن ، قطعه آزمایشی با گام زاویه ای 5 درجه برای آزمایش چرخانده شده بود . در مرحله آموزش ، 180 قطعه ماشین انتخاب شده و بصورت عمودی تنظیم شده بودند. در همین حین ، 1080 قطعه پس زمینه بصورت تصادف به عنوان منفی ها برای آموزش انتخاب شده بودند. اندازه قطعه برای این روش ها 61*31 پیکسل بود که این نپموضوع با توجه به اینکه یک بیت از اطلاعات پس زمینه به عملکرد این روش ها در شرایط طبیعی را منفعت می رساند.

شکل 14 مقایسه عملکرد بین روش ما و 3 روش دیگر را در مجموعه داده تورنتو نشان می دهد. محور افقی نرخ فراخوانی برای وسایل نقلیه بوده و محور عمودی دقت را نشان می دهد. خط آبی نشانگر نتیجه روش ما است که دقت بالاتری در مقایسه با دیگر روش ها بدست آورد. روش ویژگی SIFT بدترین عملکرد در بین 4 روش این آزمایش را داشت. زمانی که نرخ فراخوانی بیشتر از 0.6 است ، روش ما هنوز هم دقت بالای خود را حفظ کرد. با این وجود ، مقادیر دقت روش های دیگر کاهش چشمگیری زمانی که نرخ فراخوانی بیشتر از 0.6 است را نشان داد . شکل 14 عملکرد بهتر روش ما را کاملا توضیح می دهد.

جدول ۲ اندازه زیرمحدوده ها برای تشخیص در مجموعه داده تورنتو

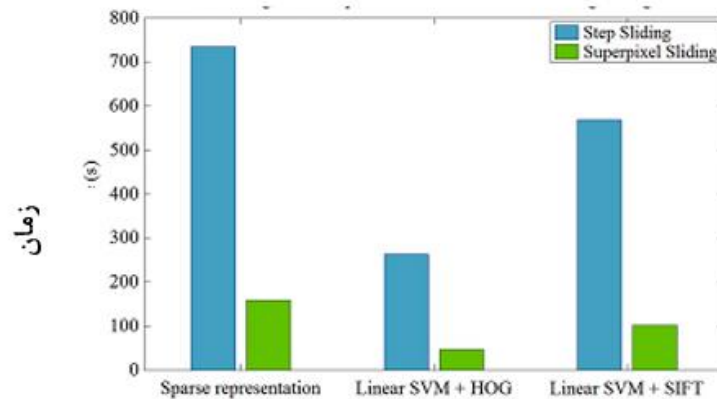
Area	1	2	3	4	5	6	7	8
Width	1644	1382	1709	2605	1537	4357	4141	2280
Length	1152	864	1358	1002	1244	1770	1932	2042

نتایج مقایسه در مجموعه داده تورنتو



شکل ۱۴. مقایسه بین روش ما و دیگر روش ها در مجموعه داده تورنتو

مقایسه زمان پردازش برای ۲ استراتژی اسکن متفاوت

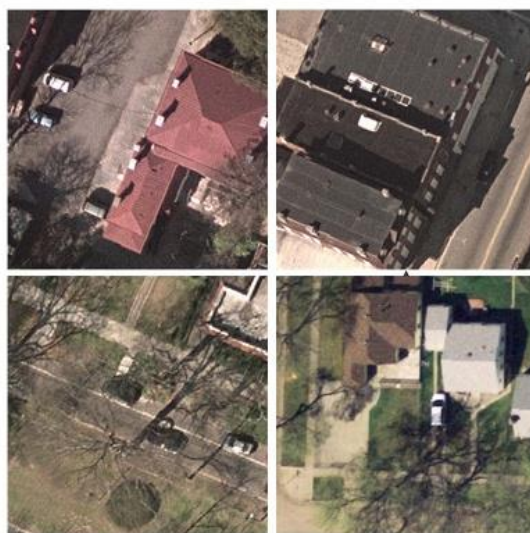


شکل ۱۵. مقایسه زمان پردازش برای ۲ استراتژی متفاوت اسکن که استراتژی اسکن با گام و لغزش ثابت و

استراتژی مبتنی بر مراکز سوپریپیکسل ها هستند

برای آزمایش اثرات استراتژی اسکن تشخیص بر اساس مراکز سوپریپیکسل ها، همچنین استراتژی لغزش سوپریپیکسل برای SVM خطی + HOG و SVM خطی + SIFT آزمایش کردیم. نتیجه آزمایش برای SVM خطی + HOG + سوپریپیکسل با خط قرمز توسط مربعات نشان داده شده است و نتیجه آزمایش برای SIFT + SVM + سوپریپیکسل با خط قرمز با دایره ها در شکل ۱۴ نشان داده شده است. نتایج تشخیص نشان می دهد که استراتژی لغزش بر اساس مراکز سوپریپیکسل ها تاثیر کمی بر دقت تشخیص در SVM خطی + HOG و SVM خطی + SIFT دارد ولی

کارایی تشخیص نسبتا بهبود یافته است همان گونه که در شکل 15 نشان داده شده است. علاوه بر این، ارائه پراکنده با استراتژی گام ثابت لغزش را بر روی این مجموعه داده آزمایش کردیم (خط آبی یا دایره ها در شکل 14 را ببینید). نتیجه نشان می دهد که روش پیشنهادی ما حتی عملکرد بهتری از ترکیب ارائه پراکنده با استراتژی گام ثابت لغزش دارد. شکل 15 مقایسه کارایی پردازش را نشان می دهد. ما آزمایشات را بر روی کامپیوتر شخصی با intel core i5-2400 CPU با 3ghz و 8-gb ram انجام دادیم. پلتفرم مورد نظر متلب بود و آزمایش تصویری با سایز 379*352 است. در شکل 15 ستون های آبی نمایان گر زمان های پردازش با استراتژی گام ثابت لغزش قدیمی را نشان می دهد. ستون های سبز زمان های پردازش استراتژی اسکن بر اساس مراکز سوپرپیکسل ها را نشان می دهد. همان گونه که در شکل 15 نشان داده شده، تمامی تشخیص ها با استراتژی گام ثابت لغزش زمان بیشتری نسبت به تشخیص متناسب با روش استراتژی اسکن بر اساس مراکز سوپرپیکسل ها مصرف می کنند. روش پیشنهادی ما مقدار کمی زیاد تر از دو روش دیگر زمان می برد. با این وجود، روش پیشنهادی ما هنوز هم کارآمد تر از تشخیص ها با استراتژی گام ثابت لغزش است.



شکل 16. 4 تصویر آزمایشی در

OIDRS

G. عملکرد OIDRS

برای تصدیق بیشتر عملکرد روش پیشنهادی ما ، OIDRS با دسترسی عموم که شامل 907 تصاویر هوایی می شود استفاده شده است. تمام تعداد وسایل نقلیه یادداشت شده در مجموعه داده حدودا 1800 است. اکثر تصاویر در این مقاله مناطق برون شهری را پوشش می دهد که منجر به تعداد زیادی از ماشین ها توسط درختان ساختمان ها و دیگر اشیا نیمه یا کام مسدود شده اند. علاوه بر این ، عوامل دیگر مانند وضوح فضایی و تغییرات مشاهده نما تشخیص ماشین را به طور منفی تحت تاثیر قرار می دهد. در آزمایش ما ، برای استفاده مستقیم دیکشنری و مدل های SVM آموزش یافته در آزمایشات قبلی ، تصاویری که وضوح فضایی متفاوت از 0.15×0.15 متر مربع داشتند بصورت دستی حذف شدند.

شکل 16 4 تصویر آزمایشی انتخابی در OIDRS را نشان می دهد. در شکل 16 ، اکثر وسایل نقلیه توسط درختان ، ساختمان ها و سایه دیگر اشیای مرتفع مسدود شده اند. در مطالعه ما ، 370 تصویر که شامل 579 وسیله نقلیه می شد برای تصدیق عملکرد روش تشخیص وسایل نقلیه پیشنهادی ما انتخاب شده اند. شکل 17 نتایج مقایسه ای روش ما و 3 روش دیگر مذکور را نشان می دهد. خط آبی با ستاره ها منحنی دقت-فراخوانی روش پیشنهادی ما را نشان می دهد. خطوط سبز، قرمز و سیاه به ترتیب منحنی دقت-فراخوانی 3 روش دیگر مذکور را نشان می دهد. در شکل 17 ، دقت تشخیص روش ما از دیگر روش ها بیشتر است. زمانی که نرخ فراخوانی بیشتر از 0.7 است ، تمامی نتایج تشخیص هر 4 روش رضایت بخش نیست. 2 دلیل برای این موضوع مسئول هستند. اول مجموعه آموزشی و مجموعه آزمایشی از مجموعه داده های مختلف حاصل شده اند که منجر به تفاوت وسایل نقلیه در دیکشنری آموزش یافته (یا مدل ها) و تصاویر آزمایشی می شود. تفاوت ها شامل اندازه ، نمای مشاهده، سطح نویز ، و وضعیت روشنایی می شوند که دست یابی به نرخ فراخوانی بیشتر در تصاویر آزمایشی را سخت تر می کند. دوم ، انسداد ها ، سایه ها و مغایرت های روشنایی در OIDRS تشخیص وسایل نقلیه را نسبتا سخت می کند که منجر به کاهش چشمگیر در دقت در حالتی که ترشلد را برای تشخیص وسایل نقلیه چالش بر انگیز با نرخ فراخوانی بالا کم میکنیم، می شود.

5. نتیجه گیری

روش جدید تشخیص وسایل نقلیه از تصاویر هوایی با وضوح بالا را ارائه کردیم که از ارائه پراکنده و بخش بندی سوپرپیکسل استفاده می کند.

توسط بخش بندی سوپرپیکسل، تصاویر هوایی ابتدا به سوپرپیکسل ها بخش بندی می شوند. بر اساس مراکز سوپرپیکسل ها، قطعات مهمی را برای آموزش و تشخیص به منظور بهره بردن از انتخاب نمونه آموزشی و کارآمد کردن اسکن تشخیص بدست آوردیم. برای تشکیل زیرمجموعه کامل و کوچک، روش انتخاب نمونه آموزشی بر اساس ارائه پراکنده پیشنهاد دادیم تا نمونه های بیشتر نماینده وار را از کل مجموعه آموزشی بزرگ انتخاب کند. با زیر مجموعه آموزشی انتخاب شده، دیکشنری ارائه پراکنده با توانایی تفکیک بالای وسایل نقلیه بدست آوردیم. کارایی الگوریتم را با استفاده از تخمین کارآمد جهت بهبود بخشیدیم تا قطعات را به حفظ جهت اصلی خود در حین آموزش و تشخیص مجبور کند (شکل 18 را ببینید).

الگوریتم خود را در 2 مجموعه داده آزمایش کردیم که شامل مجموعه داده تورنتو و OIDS می شود. چندین مدل بروز و پیشرفته (شامل +خطی SVM، SVM + HOG، SVM هسته و مرکزی، SVM + SIFT) با روش ما مقایسه شده اند. مقایسات نتایج تشخیص نشان می دهد که روش ما نتیجه تشخیص رضایت بخشی را بدست آورده و از روش های عملکرد بهتری داشته است. 3 عامل دقت تشخیص در روش ما را تحت تاثیر قرار می دهند به اسم های اندازه بخش بندی سوپرپیکسل، زمان تکرار انتخاب نمونه، و کامل بودن مجموعه آموزشی کلی اصلی. تحلیل آزمایشی در مورد زمان تکرار روش انتخاب نمونه آموزشی و کارایی پردازش نیز در آزمایشات ما ارائه شده اند.

اگرچه استراتژی اسکن کردن بر اساس سوپرپیکسل را در روش خود برای بهبود کارایی تشخیص اضافه کردیم، ارائه پراکنده همچنان پیچیدگی محاسباتی بیشتری نسبت به روش های SVM با همان استراتژی اسکن دارد.

بنابراین در مطالعات آینده، ساختار طبقه بندی سلسله مراتبی را برای بهبود بیشتر کارایی و دقت تشخیص بررسی خواهیم کرد.



شکل 18 . نتایج تشخیص وسایل نقلیه روش پیشنهادی در 2 زیرمحدوده تصویر آزمایشی. در هر دوی (a) و (b) ، مکان هایی با مربع ها مناطق تشخیص داده شده به عنوان ماشین هستند. مربع ها با رنگ قرمز نشانگر تشخیص اشتباه و مربع ها با رنگ سبز نشانگر تشخیص صحیح هستند. دوباره تشخیص داده شده ها به عنوان تشخیص اشتباه فرض شده اند

REFERENCES

- [1] B. Tian, Q. Yao, Y. Gu, K. Wang, and Y. Li, "Video processing techniques for traffic flow monitoring: A survey," in Proc. 14th Int. IEEE ITSC, 2011, pp. 1103–1108.
- [2] K. Mandal et al., "Road traffic congestion monitoring and measurement using active RFID and GSM technology," in Proc. IEEE 14th ITSC, 2011, pp. 1375–1379.
- [3] R. Du et al., "Effective urban traffic monitoring by vehicular sensor networks," IEEE Trans. Veh. Technol., vol. 64, no. 1, pp. 273–286, Jan. 2014.
- [4] S. Kamijo, Y. Matsushita, K. Ikeuchi, and M. Sakauchi, "Traffic monitoring and accident detection at intersections," IEEE Trans. Intell. Transp. Syst., vol. 1, no. 2, pp. 108–118, Jul. 2000.
- [5] W. Liu, F. Yamazaki, and T. T. Vu, "Automated vehicle extraction and speed determination from QuickBird satellite images," IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 4, no. 1, pp. 75–82, Mar. 2011.
- [6] Z. Zheng, X. Wang, G. Zhou, and L. Jiang, "Vehicle detection based on morphology from highway aerial images," in Proc. IEEE IGARSS, 2012, pp. 5997–6000.
- [7] J. Leitloff, S. Hinz, and U. Stilla, "Vehicle detection in very high resolution satellite images of city areas," IEEE Trans. Geosci. Remote Sens., vol. 48, no. 7, pp. 2795–2806, Jul. 2010.
- [8] R. Ruskoné, L. Guigues, S. Airault, and O. Jamet, "Vehicle detection on aerial images: A structural approach," in Proc. Int. Conf. Pattern Recog., 1996, pp. 900–900.
- [9] X. Jin, and C. H. Davis, "Vehicle detection from high-resolution satellite imagery using morphological shared-weight neural networks," Image Vis. Comput., vol. 25, no. 9, pp. 1422–1431, Sep. 2007.

- [10] B. Salehi, Y. Zhang, and M. Zhong, "Automatic moving vehicles information extraction from single-pass WorldView-2 imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 1, pp. 135–145, Feb. 2012.
- [11] A. Kembhavi, D. Harwood, and L. S. Davis, "Vehicle detection using partial least squares," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 6, pp. 1250–1265, Jun. 2011.
- [12] H. Grabner, T. T. Nguyen, B. Gruber, and H. Bischof, "On-line boostingbased car detection from aerial images," *ISPRS J. Photogramm. Remote Sens.*, vol. 63, no. 3, pp. 382–396, May 2008.
- [13] T. Moranduzzo and F. Melgani, "Detecting cars in UAV images with a catalog-based approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 10, pp. 6356–6367, Oct. 2014.
- [14] Z. Zheng et al., "A novel vehicle detection method with high resolution highway aerial image," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 6, pp. 2338–2343, Dec. 2013.
- [15] X. Cao, C. Wu, P. Yan, and X. Li, "Linear SVM classification using boosting HOG features for vehicle detection in low-altitude airborne videos," in *Proc. 18th IEEE ICIP*, pp. 2421–2424, 2011.
- [16] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Comput. Soc. Conf. CVPR*, 2005, pp. 886–893.
- [17] T. Moranduzzo and F. Melgani, "A SIFT-SVM method for detecting cars in UAV images," in *Proc. IEEE IGARSS*, 2012, pp. 6868–6871.
- [18] S. Maji, A. C. Berg, and J. Malik, "Classification using intersection kernel support vector machines is efficient," in *Proc. IEEE Conf. CVPR*, 2008, pp. 1–8.
- [19] J. Wright et al., "Sparse representation for computer vision and pattern recognition," *Proc. IEEE*, vol. 98, no. 6, pp. 1031–1044, Jun. 2010.
- [20] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.
- [21] A. Levinstein et al., "TurboPixels: Fast superpixels using geometric flows," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 12, pp. 2290–2297, Dec. 2009.
- [22] S. Wang, H. Lu, F. Yang, and M.-H. Yang, "Superpixel tracking," in *Proc. IEEE ICCV*, 2011, pp. 1323–1330.
- [23] X. Li, T. Jia, and H. Zhang, "Expression-insensitive 3D face recognition using sparse representation," in *Proc. Conf. CVPR*, 2009, pp. 2575–2582.
- [24] F. Chen, H. Yu, and R. Hu, "Shape sparse representation for joint object classification and segmentation," *IEEE Trans. Image Process.*, vol. 22, pp. 992–1004, Mar. 2013.
- [25] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, Dec. 2006.
- [26] J. Zepeda, C. Guillemot, and E. Kijak, "Image compression using sparse representations and the iteration-tuned and aligned dictionary," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 5, pp. 1061–1073, Sep. 2011.
- [27] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution as sparse representation of raw image patches," in *Proc. IEEE Conf. CVPR*, 2008, pp. 1–8.
- [28] J. Yang, K. Yu, Y. Gong, and T. Huang, "Linear spatial pyramid matching using sparse coding for image classification," in *Proc. IEEE Conf. CVPR*, 2009, pp. 1794–1801.
- [29] J. Yang and M.-H. Yang, "Top-down visual saliency via joint CRF and dictionary learning," in *Proc. IEEE Conf. CVPR*, 2012, pp. 2296–2303.
- [30] M. Cheng, C. Wang, and J. Li, "Sparse representation based pansharpening using trained dictionary," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 293–297, Jan. 2013.
- [31] N. Yokoya and A. Iwasaki, "Object localization based on sparse representation for remote sensing imagery," in *Proc. IEEE IGARSS*, 2014, pp. 2293–2296.
- [32] A. P. Moore, S. Prince, J. Warrell, U. Mohammed, and G. Jones, "Superpixel lattices," in *Proc. IEEE Conf. CVPR*, 2008, pp. 1–8.
- [33] R. Achanta et al., "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.

- [34] J. Wang and X. Wang, "VCells: Simple and efficient superpixels using edge-weighted centroidal Voronoi tessellations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 6, pp. 1241–1247, Jun. 2012.
- [35] M.-Y. Liu, R. Chellappa, O. Tuzel, and S. Ramalingam, "Entropy-rate clustering: Cluster analysis via maximizing a submodular function subject to a matroid constraint," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 1, pp. 99–112, Jan. 2013.
- [36] S. Hinz and A. Baumgartner, "Vehicle detection in aerial images using generic features, grouping, and context," in *Proc. Pattern Recognit.*, 2001, pp. 45–52.
- [37] H. Moon, R. Chellappa, and A. Rosenfeld, "Optimal edge-based shape detection," *IEEE Trans. Image Process.*, vol. 11, no. 11, pp. 1209–1227, Nov. 2002.
- [38] S. Hinz, "Detection and counting of cars in aerial images," in *Proc. ICIP*, 2003, vol. 2, pp. 997–1000.
- [39] P. Reinartz, M. Lachaise, E. Schmeer, T. Krauss, and H. Runge, "Traffic monitoring with serial images from airborne cameras," *ISPRS J. Photogramm. Remote Sens.*, vol. 61, no. 3/4, pp. 149–158, Dec. 2006.
- [40] J.-Y. Choi and Y.-K. Yang, "Vehicle detection from aerial images using local shape information," *Adv. Image Video Technol.*, vol. 5414, pp. 227–236, 2009.
- [41] S. M. Khan, H. Cheng, D. Matthies, and H. Sawhney, "3D model based vehicle classification in aerial imagery," in *Proc. IEEE Conf. CVPR*, 2010, pp. 1681–1687.
- [42] J. Xiao, H. Cheng, H. Sawhney, and F. Han, "Vehicle detection and tracking in wide field-of-view aerial video," in *Proc. IEEE Conf. CVPR*, 2010, pp. 679–684.
- [43] H.-Y. Cheng, C.-C. Weng, and Y.-Y. Chen, "Vehicle detection in aerial surveillance using dynamic Bayesian networks," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2152–2159, Apr. 2012.
- [44] T. Moranduzzo and F. Melgani, "Automatic car counting method for unmanned aerial vehicle images," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 3, pp. 1635–1647, Mar. 2014.
- [45] W. Shao, W. Yang, G. Liu, and J. Liu, "Car detection from high-resolution aerial imagery using multiple features," in *Proc. IEEE IGARSS*, 2012, pp. 4379–4382.
- [46] H. Wang et al., "Object detection in terrestrial laser scanning point clouds based on Hough forest," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 10, pp. 1807–1811, Oct. 2014.
- [47] H. Moon, R. Chellappa, and A. Rosenfeld, "Performance analysis of a simple vehicle detection algorithm," *Image Vis. Comput.*, vol. 20, no. 1, pp. 1–13, Jan. 2002.
- [48] X. Zhou, W. Jiang, Y. Tian, and Y. Shi, "Kernel subclass convex hull sample selection method for SVM on face recognition," *Neurocomputing*, vol. 73, no. 10/12, pp. 2234–2246, Jun. 2010.
- [49] F. Nie, H. Wang, H. Huang, and C. Ding, "Early active learning via robust representation and structured sparsity," in *Proc. 23rd Int. Joint Conf. Artif. Intell.*, 2013, pp. 1572–1578.
- [50] Z. Jiang, Z. Lin, and L. Davis, "Label consistent K-SVD: Learning a discriminative dictionary for recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 11, pp. 2651–2664, Nov. 2013.
- [51] D. L. Donoho, "For most large underdetermined systems of linear equations the minimal," *Commun. Pure Appl. Math.*, vol. 59, no. 6, pp. 797–829, Jun. 2006.
- [52] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
- [53] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
- [54] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. 8th IEEE ICCV*, 2001, pp. 416–423.
- [55] G. Chen, Y. Ding, J. Xiao, and T. X. Han, "Detection evolution with multi-order contextual co-occurrence," in *Proc. IEEE Conf. CVPR*, 2013, pp. 1798–1805. Z