

بسط دوگانه سوپر پیکسل به منظور تعیین ناحیه‌های موثر در حل مسائل بینایی ماشین مبتنی بر قطعه‌بندی

مریم تقی زاده^۱، عبدالله چاله‌چاله^۲

چکیده

یکی از روش‌های موثر برای شناسایی دیداری (شامل طبقه‌بندی، شناسایی اشیاء و برجسب‌گذاری معنایی تصویر)، تعیین ناحیه‌های محتمل بر وجود شیء بنام نامزدهای شیء (Object proposals) است. در این مقاله یک روش کارآمد با تکیه بر تعیین ناحیه‌های موثر مبتنی بر قطعه‌بندی ناحیه‌ای به نام بسط دوگانه سوپر پیکسل SDE (Superpixel Dual Extension) به منظور بهبود کیفیت ناحیه‌بندی پیشنهاد شده است. روش پیشنهادی شامل دو مرحله است. در مرحله ابتدایی با اعمال الگوریتم قطعه‌بندی پایین به بالا تصویر با تعدادی ناحیه به عنوان سوپر پیکسل نمایش داده می‌شود. سپس هر سوپر پیکسل با توجه به ۸-همسایگی و بر اساس مجموعه‌ای از شرایط تعریف شده به ناحیه‌های مجاور خود بسط داده می‌شود. از امتیازات این بسط ایجاد ناحیه‌هایی است که به خوبی می‌تواند کل یک شیء را در برگیرد. سپس با اعمال توصیفگرهایی نظیر رنگ، بافت و نقاط کلیدی و استخراج ویژگی از هر ناحیه، این ناحیه‌ها می‌توانند برای حل مسائل مختلف پردازش تصویر و بهبود کارایی آنها بکار گرفته شوند. در این مقاله، برای بررسی دقیق کیفیت ناحیه‌های به دست آمده، مجموعه‌ای از معیارهای شناخته شده قطعه‌بندی ناحیه‌ای شامل میزان همپوشانی ناحیه، مقدار فراخوانی، مساحت زیر نمودار پارامتر فراخوانی و همبستگی جفت پیکسل‌ها اندازه‌گیری شده‌اند. همچنین برای ارزیابی روش پیشنهادی، تاثیر این روش در دقت طبقه‌بندی اشیاء بر روی مجموعه داده MSRC بررسی شده است. نتایج، افزایش کیفیت و دقت را هم برای ناحیه‌های به دست آمده به میزان ۷٪ برای الگوریتم قطعه‌بندی مبتنی بر گراف و ۱۴٪ برای الگوریتم قطعه‌بندی مبتنی بر خوشه‌بندی و همچنین مقدار ۱۱٪ افزایش دقت را برای طبقه‌بندی تصویر تایید می‌کند.

کلیدواژه‌ها

بسط، سوپر پیکسل، قطعه‌بندی ناحیه‌ای، طبقه‌بندی

۱ مقدمه

نامزدهای شیء شناخته می‌شود. در این روش هدف تعیین ناحیه‌های دقیق با تعداد مناسبی است که احتمال وجود شیء در آنها بسیار زیاد است. سپس با به کارگیری شناسایی کننده‌های مناسب شیء و الگوریتم‌های یادگیری ماشین می‌توان از این نواحی برای جستجوی شیء و برجسب‌گذاری شیء استفاده کرد. در این راه-حل، تصویر را با تعدادی ناحیه نمایش می‌دهند که بر اساس الگوریتم‌های رایج قطعه‌بندی به دست می‌آیند. سپس با ارائه‌ی الگوریتم‌های مختلف تعدادی ناحیه موثر با تکیه بر همان قطعه‌بندی اولیه به دست آورده و سپس با به کارگیری توصیف‌گرها و شناسایی کننده‌های مناسب شیء به شناسایی اشیاء می‌پردازند. از

در دهه‌ی اخیر یکی از روش‌های بسیار محبوب و موثر برای حل مساله‌ی شناسایی دیداری گروه‌بندی مفهومی است که به‌عنوان

این مقاله در بهمن‌ماه ۱۳۹۶ دریافت، در شهریورماه ۱۳۹۷ بازنگری و در همان ماه پذیرفته شد.

^۱ دانشجوی دکتری معماری کامپیوتر، گروه مهندسی کامپیوتر، دانشگاه رازی
رایانامه: taghzadehmail@gmail.com

^۲ دانشکده فنی مهندسی، گروه مهندسی کامپیوتر، دانشگاه رازی

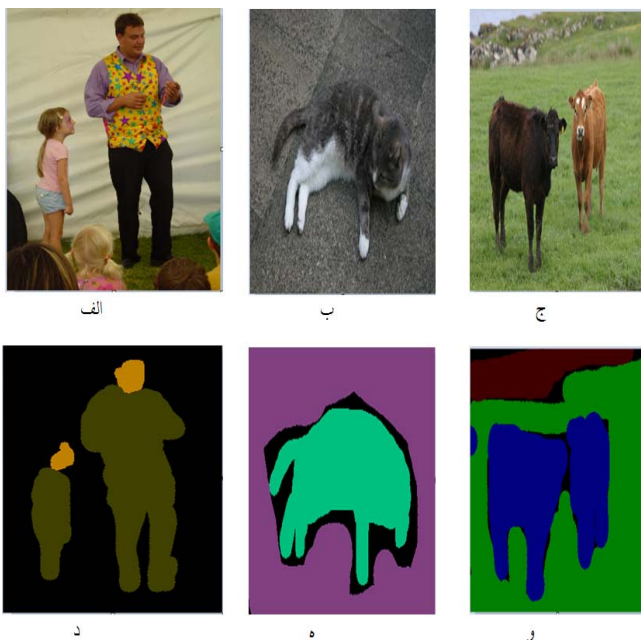
رایانامه: Chalechale@razi.ac.ir

ج). روش پیشنهادی SDE با توجه به این ویژگی‌ها شامل دو حالت متفاوت است که مشخصه‌های کلی تصاویر را تا حدود زیادی پوشش می‌دهد. هدف اصلی ما در این مقاله پیشنهاد یک الگوریتم بسط بر اساس ۸-همسایگی برای سوپرپیکسل‌های تصویر، به منظور ایجاد ناحیه‌های دقیق منطبق بر اشیای تصویر می‌باشد که ضعف الگوریتم‌های قطعه‌بندی را کاهش می‌دهد. برای ارزیابی، کیفیت ناحیه‌های به دست آمده بطور کامل محاسبه و بررسی شده‌اند. همچنین روش پیشنهادی را برای حل مساله‌ی طبقه‌بندی تصویر بررسی کرده‌ایم. بطور خلاصه مشارکت اصلی این مقاله عبارت است از:

(۱) ارائه‌ی یک الگوریتم دوگانه برای بسط ناحیه به منظور افزایش کارایی

(۲) افزایش دقت الگوریتم‌های قطعه‌بندی مختلف

(۳) افزایش دقت طبقه‌بندی تصویر با الگوریتم پیشنهادی



شکل ۱ مجموعه‌ای از تصاویر MSRC به همراه تصاویر مرجع درستی^۱ قطعه‌بندی شده آن‌ها [۲۲]. (الف) در این تصویر بخش‌های مختلف شیء در رنگ و بافت متفاوت هستند، (ب) در این تصویر شیء دارای بافت مجزا اما رنگ یکسان با پیش زمینه است، (ج) در این تصویر، اشیاء بافت یکسان و رنگ متفاوت دارند. (د) تصویر قطعه‌بندی شده تصویر الف، (ه) تصویر قطعه‌بندی شده تصویر ب، (و) تصویر قطعه‌بندی شده تصویر ج.

در ادامه ساختار مقاله بدین صورت است: در بخش دوم مروری بر کارهای پیشین در تعیین ناحیه‌های تصویر با هدف شناسایی دیداری تصویر انجام شده است. در بخش بعدی روش پیشنهادی معرفی و جزئیات آن بیان شده است. بخش چهارم ارزیابی کاملی

اولین تحقیقاتی که شناسایی را بر پایه‌ی ناحیه پیشنهاد داد می‌توان به [۱] اشاره نمود. در ادامه [۲] و [۳] با به کارگیری ناحیه‌بندی به شناسایی اشیاء و قطعه‌بندی معنایی پرداختند.

در تعیین نامزدهای شیء تعداد ناحیه‌های کمتر و با دقت بیشتر یکی از اهداف اصلی و چالش‌های مهم می‌باشد. هر چه تعداد ناحیه‌ها کمتر باشد زمان پردازش کوتاه‌تر و قابل قبول‌تر خواهد شد و برای مسائل مختلف عملی و همچنین برای کاربردهای بی‌درنگ نیز می‌تواند کارآمد و مناسب باشد. بنابراین، افزایش دقت نامزدهای محتمل بر وجود اشیاء، دقت مسائل شناسایی دیداری را در مجموع به‌طور چشمگیری افزایش خواهد داد. در روش‌های اولیه برای تعیین ناحیه‌های مبتنی بر وجود شیء می‌توان به اسکن کردن پنجره‌های لغزنده^۱ در تمام ابعاد تصویر اشاره نمود که در مجموع فضای جستجو در این روش به اندازه‌ای زیاد است که می‌توان این روش را از نگاه محاسباتی به نوعی غیرممکن فرض کرد [۴] و [۵]. بنابراین محققان سراغ ارائه‌ی روش‌های جستجوی ابتکاری با هدف کاهش فضای جستجو رفتند. در کارهای گزارش شده اخیر تعداد ناحیه‌های استخراج شده از تصاویر به طور متوسط به ۱۰۰۰۰ تا ۲۰۰۰۰ کاهش یافت [۶] تا [۸]. در حالت کلی این روش‌ها بر اساس دو گروه متفاوت «وابسته به اشیاء خاص»^۲ یا «مستقل از دسته خاص»^۳ می‌باشند. در این مقاله به دنبال ارائه‌ی روشی مستقل از اشیاء هستیم که ناحیه‌های استخراج شده از تصویر را با به کارگیری محاسبات غیرپیچیده و اندک به نواحی با دقت مناسب‌تری برای شناسایی شیء تبدیل نماید. این روش با نگاه به افزایش توان الگوریتم‌های قطعه‌بندی پیشنهاد شده است. استخراج دقیق ناحیه‌های محتمل بر وجود شیء در تصویر در کاربردهای مختلف و به ویژه در اعمال الگوریتم‌های یادگیری چندمنونه‌ای و یادگیری عمیق که استخراج ناحیه‌های موثر اهمیت زیادی در بهبود کارایی آن‌ها و کاهش خطا دارد، تاثیر بسزایی خواهد داشت.

این مقاله یک روش موثر با توجه به نمایش تصویر با سوپرپیکسل‌ها معرفی می‌نماید که مرحله بسط سوپرپیکسل را بر اساس ۸-همسایگی هر سوپرپیکسل به دو شکل مختلف انجام می‌دهد. ناحیه‌های پایانی به دست آمده دارای کیفیت مناسب و تاثیرگذار بر بهبود دقت در مسائل بینایی ماشین هستند. روش پیشنهادی با نام بسط دوگانه سوپرپیکسل SDE معرفی شده است. این الگوریتم بر اساس سه ویژگی مهم از تصاویر شکل گرفته است. یک شیء می‌تواند از بخش‌های مختلفی تشکیل شده باشد که دارای رنگ و بافت متفاوتی با یکدیگر باشند (شکل ۱- الف). یک شیء می‌تواند رنگ یکسان با محیط احاطه شده‌ی اطراف خود داشته اما بافت آن‌ها متفاوت باشند (شکل ۱- ب). یک شیء می‌تواند دارای رنگ‌های متفاوت ولی بافت یکسان باشد (شکل ۱-

¹ Sliding window

² Class specific

³ Class independent

تاریخچه‌ی الگوریتم‌های قطعه‌بندی به یک بازه‌ی زمانی ۷۰ ساله برمی‌گردد که همچنان مساله‌ی بازی است و به‌طور کامل حل نشده است و محققان زیادی به ارائه‌ی روش برای آن پرداخته‌اند. نکته‌ی مهم دیگری که پرداختن به این موضوع را جذاب‌تر و مهم‌تر نموده است به‌کارگیری قطعه‌بندی در حل مسائل مختلف شناسایی دیداری است [۱۵]. بنابراین ما به دنبال روشی مناسب برای بهبود ضعف قطعه‌بندی مبتنی بر ناحیه‌هستیم که در پایان قادر به تعیین ناحیه‌های مناسب به‌منظور به‌کارگیری در مسائل مختلف شناسایی دیداری باشد. روش‌های مختلفی تاکنون برای مساله‌ی قطعه‌بندی ارائه شده است. در [۱۶] مروری بر روش‌های مختلف قطعه‌بندی انجام شده است. بطور کلی روش قطعه‌بندی مبتنی بر الگوریتم‌های پایین به بالا^۶، بالا به پایین^۷ و یا ترکیبی می‌تواند باشد و روش‌های متعددی در این سه چارچوب معرفی شده‌اند. از مجموع این روش‌ها می‌توان به برش نرمالیزه^۸ [۱۷]، جابجایی سریع^۹ [۱۸]، جابجایی میانگین^{۱۰} [۱۹]، مبتنی بر گراف^{۱۱} (EGBS) [۲۰] و خوشه‌بندی تکراری خطی ساده^{۱۲} (SLIC) [۲۱] اشاره نمود. الگوریتم EGBS با به‌کارگیری مرز میان ناحیه‌ها تصویر را قطعه‌بندی می‌کند. در روش‌های زیادی که برای شناسایی اشیاء ارائه شده است این الگوریتم به‌عنوان اولین مرحله به‌کار گرفته شده است [۶] و [۷]. روش SLIC بر اساس اجرای یک خوشه‌بندی تکراری خطی است. این روش خوشه‌بندی پیکسل‌ها را در یک فضای ۵ بعدی توسط مقادیر L ، a و b در فضای رنگی CIELAB انجام می‌دهد و مختصات x و y پیکسل را نیز در نظر می‌گیرد. این روش از اندازه‌گیری فاصله برای فشرده‌سازی و منظم بودن شکل سوپریکسل‌ها استفاده می‌کند. یکی از امتیازات این روش کنترل کردن تعداد سوپریکسل‌های استخراج شده است.

با توجه به اینکه نتایج به‌دست آمده از قطعه‌بندی EGBS و SLIC نتایج خوبی را گزارش داده‌اند، ما نیز در این مقاله این دو روش را برای مرحله‌ی ابتدایی به‌کار برده‌ایم. سپس یک الگوریتم بسط دوگانه مناسب به‌منظور تعیین کاندیدهای قطعه‌بندی شده مناسب با افزایش دقت قطعه‌بندی پیشنهاد شده که در دیگر مسائل بینایی ماشین نیز قابل استفاده است. در بخش بعدی تشریح این الگوریتم آورده شده است.

۳ الگوریتم بسط دوگانه سوپریکسل

برای استخراج ناحیه‌های دقیق و مبتنی بر اشیای مختلف، روش مناسبی با عنوان بسط دوگانه سوپریکسل (SDE) معرفی می‌کنیم. این روش را می‌توان در کاربردهای مختلف از پردازش تصویر و

از روش پیشنهادی به همراه نتایج به‌دست آمده ارائه می‌نماید. بحث و تحلیل روش پیشنهادی در بخش پنجم و سپس نتیجه‌گیری در بخش پایانی گزارش شده است.

۲ پیش زمینه

گروه‌بندی ادراکی^۱ یک تصویر به عنوان یک مرحله‌ی ابتدایی پیش پردازش در بسیاری از کاربردهای پردازش تصویر مانند بازیابی معنایی مبتنی بر محتوا [۹]، آنالیز تصویر [۱۰] و موارد دیگر به‌کار گرفته شده است. علاوه بر این، در مساله‌ی شناسایی شیء یکی از روش‌هایی که در دهه‌ی اخیر توجه محققان را به خود جلب نموده، نامزدهای شیء است که مبتنی بر گروه‌بندی ادراکی است. این روش یا وابسته به دسته خاصی از اشیاء می‌باشد یا مستقل از شیء است و خروجی‌های آن در مجموع می‌تواند در دو حالت متفاوت جعبه‌های مرزی^۲ یا همان پنجره‌های تصویری مانند Objectness [۷] و جستجوی انتخابی [۸] یا کاندیدهای قطعه‌بندی شده^۳ مانند MCG^۴ [۶] و CPMC^۵ باشد [۱۱]. امروزه روش نامزدهای شیء به‌عنوان مرحله‌ی پایه برای شناسایی شیء و برچسب‌گذاری تصویر در روش‌های کلاسیک استخراج ویژگی و روش‌های پیشرفته استخراج ویژگی که مبتنی بر یادگیری عمیق هستند، بکار گرفته می‌شود [۱۲] تا [۱۴]. به‌طور کلی در روش‌های پیشنهاد شده اخیر، به دنبال کاهش فضای جستجو و ارائه‌ی ناحیه‌های مناسبی هستند که اشیاء را بطور دقیق احاطه نمایند. روش Objectness با در نظر گرفتن چندین فرضیه، مکان‌هایی از تصویر را به‌صورت پنجره‌هایی در نظر گرفته و آن‌ها را به‌عنوان کاندیدهای اشیاء فرض می‌کند. در روش جستجوی انتخابی با تکیه بر مجموعه‌ای از ویژگی‌های مکمل در تصویر و به‌کارگیری ساختار قطعه‌بندی سلسله‌مراتبی ناحیه‌هایی به‌عنوان کاندیدهای اشیاء انتخاب می‌شود. روش MCG یک رویکرد قطعه‌بندی سلسله‌مراتبی چندمقیاسی پیشنهاد می‌کند که مبتنی بر ارائه‌ی الگوریتم قطعه‌بندی کارآمدی است. برای کاهش تعداد نواحی محتمل بر اشیاء، یک الگوریتم ترکیب‌بندی جفتی، سه تایی و چهارتایی از ناحیه پیشنهاد داده شده است که این ترکیب در چندمقیاس و به‌صورت سلسله‌مراتبی انجام می‌شود. این ترکیب‌سازی با هدف کاهش تعداد ناحیه انجام می‌شود. در این مقاله به دنبال راه‌حل کارآمد و مناسبی هستیم که بر اساس همسایگی ناحیه‌های یک تصویر، بسط را انجام دهد. این روش پیشنهادی به‌طور مستقیم ایراد قطعه‌بندی را نیز بهبود داده و ناحیه‌های جدید به‌دست آمده می‌توانند برای مسائل مختلف در بینایی ماشین به‌کار گرفته شوند.

هدفی که در تمامی روش‌های ارائه شده در بالا وجود دارد، حذف کاستی‌های موجود در الگوریتم‌های قطعه‌بندی است.

⁶ Bottom-up

⁷ Top-down

⁸ Normalized-Cut

⁹ Quick shift

¹⁰ Mean shift

¹¹ Efficient graph-based segmentation

¹² Simple linear iterative clustering

¹ Perceptual grouping

² Bounding box

³ Segmented candidate

⁴ Multiscale combinatorial grouping

⁵ Constrained parametric min-cuts

این مقاله برای مرحله‌ی اول از دو روش مختلف EGBS و SLIC استفاده نموده و سپس در بخش ارزیابی تاثیر روش پیشنهادی بر روی هر دو نوع الگوریتم بررسی خواهد شد.

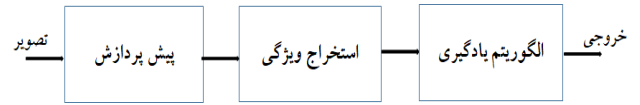
در روش EGBS ناحیه‌های تصویر مبتنی بر پارتیشن‌بندی گراف به دست می‌آیند. هر پیکسل تصویر یک گره از گراف می‌باشد و در گراف از ساختار ۴- همسایگی یا ۸- همسایگی استفاده می‌شود. برای لبه‌های گراف یک مقدار عددی تخصیص داده شده که مبتنی بر معیار عدم شباهت است و ویژگی به کار گرفته شده مقدار رنگ پیکسل است. در واقع مقدار اختلاف رنگ دو پیکسل مجاور محاسبه می‌شود. دو پیکسل مجاور در صورتی در یک سوپرپیکسل قرار گرفته می‌شوند که مقدار عددی نسبت داده شده به لبه گراف از یک مقدار از پیش تعیین شده کمتر باشد. این روند زمانی متوقف می‌شود که مقدار عددی لبه از مقدار تعیین شده بیشتر شود و در پایان این ترکیب‌سازی یک سوپرپیکسل به دست می‌آید. این مراحل تا زمانی که همه پیکسل‌های تصویر ناحیه‌بندی شوند ادامه پیدا می‌کند.

در روش SLIC، سوپرپیکسل‌ها بر اساس به کارگیری خوشه‌بندی kmeans بر روی ویژگی‌ها به دست می‌آیند و ویژگی به کار گرفته شده تشابه رنگ پیکسل در فضای رنگی CIELAB می‌باشد. همچنین مختصات سطر و ستون پیکسل به عنوان ویژگی مکانی استفاده می‌شود. در مجموع بردار ویژگی آن به طول ۵ تایی و بر اساس یک ترکیب خطی از ویژگی رنگ و ویژگی مکان به دست می‌آید.

سپس در مرحله‌ی دوم الگوریتم بسط دوگانه ارائه گردیده است که هر ناحیه سوپرپیکسل را بر اساس همسایگی خود بسط می‌دهد. این الگوریتم پیشنهادی در دو گونه متفاوت «حالت ثابت»- fixed mode و «حالت همه»- all-mode پیشنهاد داده شده است. الگوریتم پیشنهادی در شکل ۳ نشان داده شده است. در ادامه به معرفی این الگوریتم و گونه‌های آن می‌پردازیم.

در «حالت ثابت»، در ابتدا برای هر سوپرپیکسل بر اساس ۸- همسایگی تمام ناحیه‌های مجاور آن در نظر گرفته می‌شود. منظور از تمام ناحیه‌های مجاور، سوپرپیکسل‌هایی هستند که با مرز سوپرپیکسل جاری حداقل در یک پیکسل مشترک باشند. در حقیقت تمام پیکسل‌هایی که بر روی مرز سوپرپیکسل قرار گرفته‌اند را در نظر گرفته و برای هر کدام به تعداد ثابت P پیکسل متوالی از ناحیه‌ی همسایه‌ی آن در خارج از مرز را به عنوان ناحیه‌ی بسط‌یافته انتخاب نموده و به سوپرپیکسل جاری اضافه می‌شوند. این مرحله برای تمامی پیکسل‌های روی مرز هر سوپرپیکسل اجرا می‌گردد. هر پیکسل تصویر می‌تواند در ۸ جهت مختلف همسایگی داشته باشد که با توجه به تعریف پیشنهاد شده در این الگوریتم، برای نقاط روی مرز سوپرپیکسل حداکثر می‌توان ۵ جهت متفاوت بیرون از سوپرپیکسل در نظر گرفت (شکل ۴ چگونگی انتخاب P پیکسل برای چند نقطه از مرز سوپرپیکسل را نشان می‌دهد). حال با توجه به اینکه هر پیکسل بر اساس جهت

بینایی ماشین اعمال نمود. به طور کلی برای حل مسائل مختلف در زمینه‌ی پردازش تصویر مبتنی بر یادگیری یک معماری مانند شکل ۲ وجود دارد.



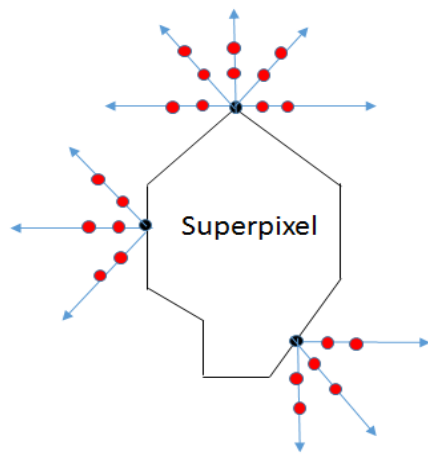
شکل ۲ معماری کلی حل مسائل بینایی ماشین

یکی از روش‌های پیش پردازش و اولیه که در حل بسیاری از مسائل پردازش تصویر و بینایی ماشین به کار گرفته می‌شود، قطعه‌بندی است. قطعه‌بندی، تصویر را به تعدادی ناحیه که در ویژگی‌هایی مانند رنگ، بافت یا مرز شیء مشابه هستند تقسیم می‌کند. یکی از روش‌های رایج قطعه‌بندی برای استخراج ناحیه‌های تصویر، الگوریتم پایین به بالا است که تمرکز آن بر روی جداسازی اشیاء نیست و در حقیقت، ایده اصلی گروه‌بندی پیکسل‌های نزدیک به هم بر اساس شباهت و همگنی در فضای ویژگی است. شباهت پیکسل‌ها بر اساس استخراج ویژگی‌های سطح پایین و سطح میانی مانند رنگ، بافت، انحناء خمیدگی و اجرای خوشه‌بندی این ویژگی‌ها انجام می‌شود. نتیجه‌ی الگوریتم قطعه‌بندی، نمایش تصویر با تعدادی ناحیه است که ناحیه‌ها به عنوان قطعه یا سوپرپیکسل در نظر گرفته می‌شوند. با توجه به اینکه در این مقاله از دو روش متفاوت تولید سوپرپیکسل استفاده شده است در ادامه ویژگی‌های به کار گرفته شده برای ایجاد سوپرپیکسل‌ها و جزئیات هر کدام آورده شده است.

ادغام کردن سوپرپیکسل‌ها نیز ایده موثر دیگری از الگوریتم‌های قطعه‌بندی است که بر اساس قوانین متفاوتی قابل اعمال است. این ترکیب‌سازی می‌تواند بر اساس شباهت و یا همسایگی و یا ترکیبی از هر دو انجام گردد. در حالت کلی، اشیاء مختلف با داشتن ویژگی‌های متفاوت و متنوع، دارای دو مشخصه بدیهی هستند. اول اینکه اشیاء از مجموعه‌ای از پیکسل‌های مجاور هم ایجاد می‌شوند. دوم اینکه، یک شیء شامل اجزای متفاوت و متنوعی بر اساس رنگ، بافت و ساختار شکل می‌باشد. از این رو، ترکیب کردن سوپرپیکسل‌های مجاور می‌تواند یک ناحیه بزرگتر و واحدی را به دست آورد که احاطه کننده کامل یک شیء باشد. بنابراین قطعه‌ی جدید به دست آمده می‌تواند نمایش بهتر و کامل‌تری از اشیاء درون تصویر باشد. اگر ناحیه‌های اجزای تصویر دقیق‌تر و کامل‌تر استخراج شوند می‌تواند تاثیر به‌سزایی بر روی مسائل مختلفی از بینایی ماشین داشته باشد.

از این رو می‌توان یک بلاک با عنوان بسط ناحیه و با تکیه بر ویژگی‌های خاص و سناریوهای مختلف در حل این گونه مسائل طراحی و در معماری مطرح شده اضافه نمود. بلاک پیشنهادی این مقاله به نام SDE است که برای گام پیش پردازش دو مرحله معرفی می‌کند. اولین مرحله تبدیل کردن تصویر به تعدادی سوپرپیکسل است که بر اساس الگوریتم‌های گروه‌بندی پایین به بالا محاسبه می‌گردد. در

و رنگ مشابه یا متفاوت باشد. بنابراین با بسط هر سوپریکسل به ناحیه‌های مجاور می‌توان پنجره‌ها یا قطعه‌هایی تولید کرد که اشیای مختلف تصویر را به طور کامل‌تری احاطه خواهند کرد. در حقیقت الگوریتم‌های قطعه‌بندی در حالت اولیه خود قادر به ایجاد ناحیه‌های صحیح و دقیق نیستند و بسط دادن ناحیه‌ها می‌تواند کارایی قطعه‌بندی را افزایش دهد. در مجموع می‌توان ایجاد ناحیه‌های پایانی در این الگوریتم را در مرحله اول بر اساس ترکیب پیکسل‌ها مبتنی بر ویژگی شباهت عنوان کرد. در ادامه در مرحله دوم، ترکیب‌سازی سوپریکسل‌های مجاور و ایجاد یک سوپریکسل بزرگتر که بر اساس همسایگی انجام می‌شود، می‌تواند به طور ضمنی هم شباهت میان سوپریکسل‌ها و هم عدم شباهت میان سوپریکسل‌ها را در نظر بگیرد. در نتیجه به طور غیرمستقیم این الگوریتم هم از ویژگی شباهت و عدم شباهت میان پیکسل‌های مجاور بهره می‌برد. در بخش ارزیابی نشان خواهیم داد که این ایده می‌تواند تاثیر فوق العاده‌ای در نتایج مسائل کاربردی پردازش تصویر داشته باشد.



شکل ۴ چگونگی بسط یک سوپریکسل بر اساس ۸-همسایگی در حالت fixed-mode با $p=2$



شکل ۵ اجرای قطعه بندی EGBS بر روی یک تصویر نمونه از مجموعه داده MSRC [۲۲]

قرار گرفتن آن بر روی مرز، در جهت بیرون از ناحیه فقط در چند جهت پیکسل مجاور دارد؛ در همان تعداد جهت مقدار ثابتی از پیکسل‌های مجاور آن به صورت خط مستقیم در نظر گرفته می‌شود. این مقدار ثابت با P نشان داده شده است. مقدار P را می‌توان بر اساس ضابطه از قبل تعریف شده مقداردهی نمود یا اینکه به صورت تصادفی/تجربی یک مقدار عددی دلخواه برای آن در نظر گرفته شود. پیشنهاد انتخاب مولفه P ثابت در این الگوریتم برای کاهش برخی از کاستی‌های الگوریتم‌های قطعه‌بندی مانند شرایطی که الگوریتم قادر به تشخیص صحیح مرز میان اشیاء متفاوت و مجاور هم در یک تصویر نیست، می‌باشد. در این شرایط در واقع بعد از اجرای الگوریتم قطعه‌بندی در مرحله اول ناحیه‌هایی از تصویر استخراج شده‌اند که هر ناحیه ممکن است بیش از یک شیء را در نظر گرفته باشد. حال با انتخاب فقط بخشی از پیکسل‌های مجاور یک ناحیه می‌توان این نقصان را تا حدودی کاهش داد. یک تصویر نمونه از اجرای الگوریتم قطعه‌بندی EGBS در شکل ۵ نمایش داده شده است. این تصویر قطعه‌بندی شده استخراج ناحیه‌های نادرست با دقت پایین را نشان می‌دهد.

Algorithm 1: (fixed-mode)

```

For each regionj ( $j \in \{1, \dots, M\}$ ),  $M$  is the number of extracted regions in the image
    PixelB=Select pixels on the region-boundary
    If PixelB have adjacent pixels out of region; then
        // Select pixels out of the region;
        Consecutively add  $P$  pixels of each external direction as adjacent pixels
    End if
End for j

```

Algorithm 2: (all-mode)

```

For each regionj ( $j \in \{1, \dots, M\}$ ),  $M$  is the number of extracted regions in the image
    Select adjacent superpixels based on an 8-neighborhood
    Add the pixels which are the out of regionj and inside adjacent superpixels
End for j

```

شکل ۳ الگوریتم پیشنهادی SDE در دو حالت «fixed-mode» و «all-mode»

در «حالت همه»، برای هر سوپریکسل تمام ناحیه‌های همسایه بر اساس ۸-همسایگی در نظر گرفته می‌شود. سپس همه‌ی پیکسل‌های ناحیه‌های مجاور برای بسط ناحیه در نظر گرفته شده و از کنار هم قرار دادن آن‌ها یک ناحیه بزرگتر نتیجه می‌شود. حال این ناحیه بزرگتر با احتمال بیشتری می‌تواند تمام اجزای شیء درون تصویر را احاطه کند. شکل ۶ یک مثال از ناحیه‌های مجاور یک سوپریکسل در این حالت بسط را نشان می‌دهد. هر کدام از این حالت‌های پیشنهادی مبتنی بر ویژگی‌های ذاتی مختلفی است که در اشیاء وجود دارد. هر شیء معمولاً از بخش‌های متفاوتی که در مجاور هم هستند تشکیل شده است. هر شیء می‌تواند دارای بافت

الگوریتم قطعه‌بندی SLIC با تعداد نواحی مختلف اجرا و مورد ارزیابی قرار گرفته است.

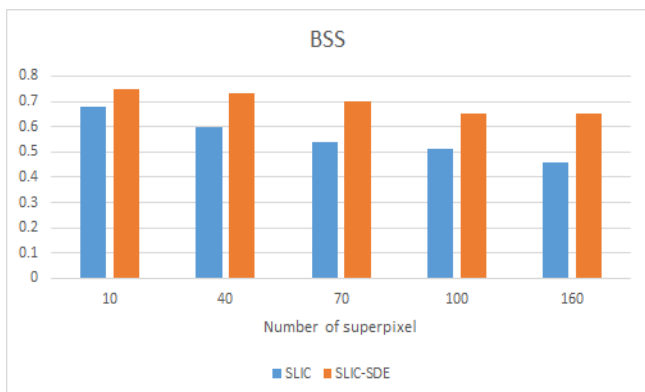
یکی از پارامترهای ارزیابی کیفیت ناحیه‌های قطعه‌بندی، معیار «بیشترین مقدار همپوشانی قطعه»^۲ (BSS) می‌باشد [۲] که ما نیز در این مطالعه BSS را برای ارزیابی ناحیه محاسبه می‌کنیم. این پارامتر نشان‌دهنده بهترین ناحیه‌ی قطعه‌بندی شده است که بیشترین همپوشانی را با شیء داشته باشد. رابطه‌ی ۱ چگونگی محاسبه‌ی مقدار همپوشانی ov را برای دو ناحیه i و j بیان می‌کند [۲]:

$$ov(i, j) = \frac{R_i \cap R_j}{R_i \cup R_j} \quad (1)$$

پارامترهای R_i و R_j سطح نواحی i و j را بیان می‌کند. در نتیجه مقدار بیشینه‌ی ov معادل پارامتر BSS خواهد بود. برای هر شیء مقدار BSS را برای الگوریتم قطعه‌بندی و الگوریتم پیشنهادی محاسبه کرده و میانگین آن را در مجموعه‌ی همه‌ی اشیاء محاسبه می‌کنیم. رابطه‌ی ۲ چگونگی محاسبه‌ی آن را نشان می‌دهد [۲]:

$$BSS = \frac{1}{N} \sum_{i=1}^N BSS_{class}(i) \quad (2)$$

تعداد دسته اشیاء برابر با N است. نتایج به دست آمده در شکل ۷ نشان داده شده است. در این شکل مقدار میانگین BSS با در نظر گرفتن تعداد ناحیه‌های مختلف در الگوریتم قطعه‌بندی SLIC محاسبه شده است. همانگونه که از شکل مشخص است بعد از اعمال الگوریتم پیشنهادی SDE به الگوریتم قطعه‌بندی SLIC (که برای راحتی آن را با SLIC-SDE نشان می‌دهیم) مقدار این پارامتر به طور متوسط ۱۴٪ افزایش می‌یابد. در حقیقت الگوریتم پیشنهادی ضعف موجود در الگوریتم قطعه‌بندی را به میزان خوبی کاهش می‌دهد. مقدار BSS برای الگوریتم قطعه‌بندی EGBS و اعمال الگوریتم پیشنهادی به آن (که برای راحتی و سادگی آن را EGBS-SDE نامگذاری می‌کنیم) محاسبه و در شکل ۸ نشان داده شده است. این شکل نیز با بهبود همپوشانی به میزان ۷٪ بعد از اعمال الگوریتم SDE، تاثیرگذاری روش پیشنهادی را تایید می‌کند.



شکل ۷ پارامتر BSS برای الگوریتم SLIC (با تعداد ناحیه‌های مختلف) و الگوریتم SLIC-SDE



شکل ۶ ناحیه‌های مجاور یک سوپرپیکسل و چگونگی بسط ناحیه در الگوریتم SDE در حالت all-mode برای بخشی از یک تصویر نمونه از مجموعه داده MSRC [۲۲]

۴ نتایج آزمایش‌ها

برای ارزیابی کامل مدل پیشنهادی دو نوع آزمایش انجام شده است. در اولین ارزیابی، کیفیت ناحیه‌های به دست آمده از اعمال الگوریتم بسط دوگانه بررسی شده است که در ادامه به تشریح آن‌ها می‌پردازیم. در دومین ارزیابی مساله‌ی طبقه‌بندی تصویر بررسی شده است. برای طبقه‌بندی تصویر هدف بررسی تاثیر روش بسط دوگانه SDE پیشنهاد شده بر روی الگوریتم یادگیری چند نمونه‌ای است. برای هر دو نوع آزمایش از مجموعه داده MSRC استفاده شده که شامل ۲۱ دسته شیء مختلف است و از اولین مجموعه داده‌ای است که برای مسائل قطعه‌بندی فراهم شده است [۲۲]. تعداد تصاویر آن ۵۹۱ و هر تصویر به طور متوسط دارای ۳ برچسب متفاوت است. برچسب‌ها هم شامل شیء و هم شامل مفاهیم^۱ هستند. منظور از مفاهیم قسمت‌هایی از تصویر مانند آسمان، آب و جاده است که دارای اشکال معنادار مانند اشیاء نیستند.

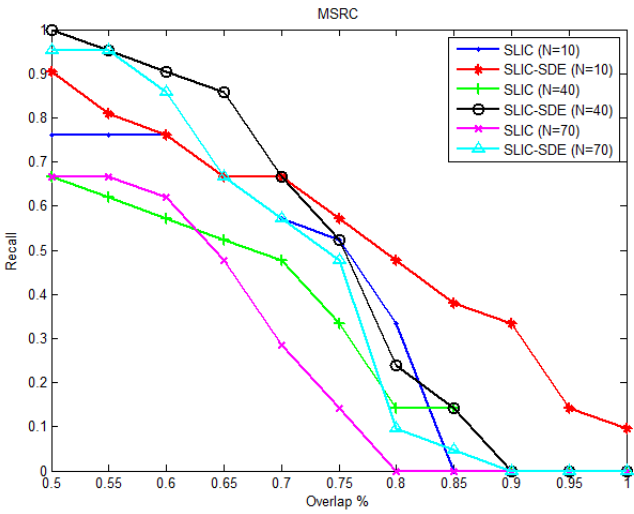
۴-۱ بررسی کیفیت ناحیه‌های مبتنی بر شیء

برای ارزیابی الگوریتم‌های قطعه‌بندی معیارهای مختلفی تاکنون ارائه شده است که به طور کلی به دو گروه ارزیابی مبتنی بر مرز و مبتنی بر ناحیه تقسیم‌بندی می‌شوند. از دیدگاه ناحیه‌بندی تصویر، تحلیل و ارزیابی را می‌توان بر اساس طبقه‌بندی جفت پیکسل‌ها و یا خوشه‌بندی مجموعه پیکسل‌ها انجام داد [۲۳] و [۲۴]. در این مقاله با توجه به اهمیت بیشتر نواحی نسبت به مرز، از پارامترهای مبتنی بر ناحیه استفاده شده است. برای بررسی کیفیت ناحیه‌های یک تصویر دو نوع الگوریتم قطعه‌بندی SLIC و EGBS و الگوریتم بسط ناحیه ارزیابی شده‌اند. نکته‌ی دیگری که در این ارزیابی‌ها فرض شده است بررسی تعداد ناحیه‌های مختلف و تاثیرگذاری آن در دیگر پارامترهای قطعه‌بندی است. از این رو،

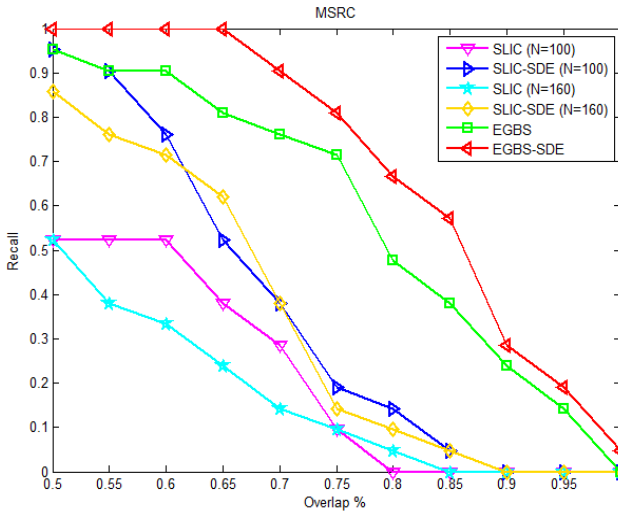
² Best overlap segmentation score

¹ Stuff

در ادامه نمودار مربوط به پارامتر Recall به ازای مقادیر مختلف آستانه‌گذاری بر روی BSS مبتنی بر الگوریتم‌های SLIC و EGBS در شکل ۱۰ ترسیم شده است. برای الگوریتم SLIC تعداد ناحیه‌های متفاوت به کار گرفته شده است.



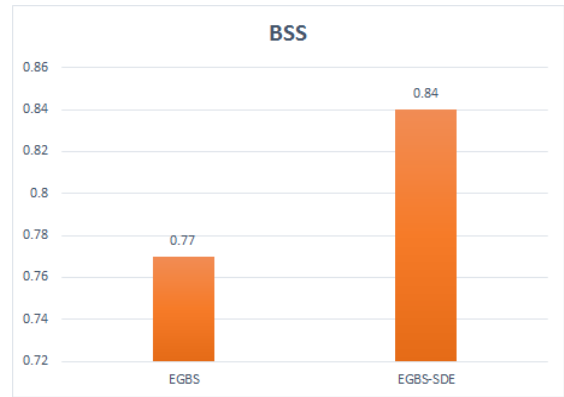
(الف)



(ب)

شکل ۱۰ نمودار Recall بر اساس مقدار همپوشانی. (الف) مقدار Recall قطعه‌بندی SLIC و SLIC-SDE با تعداد ناحیه‌های ۱۰، ۴۰ و ۷۰. (ب) مقدار Recall برای SLIC-SDE، SLIC با ۱۰۰ و ۱۶۰ ناحیه) و EGBS و EGBS-SDE. نتایج کیفیت ناحیه‌های به‌دست آمده با استفاده از الگوریتم SDE را نشان می‌دهد.

یکی دیگر از پارامترهای ارزیابی مقدار مساحت زیرنمودار^۱ (AUC) است که به‌عنوان یک عدد به کار گرفته می‌شود و نشان دهنده‌ی کارایی الگوریتم است [۲۵]. در ادامه مقدار AUC برای نمودار ارائه شده Recall مبتنی بر BSS که در شکل ۱۰ نشان داده شده، محاسبه و در شکل‌های ۱۱ و ۱۲ آورده شده است. با توجه به نمودار، افزایش AUC بعد از اعمال SDE به الگوریتم‌های قطعه‌بندی به وضوح دیده می‌شود.



شکل ۱۱ پارامتر BSS در الگوریتم EGBS و EGBS-SDE

علاوه بر این در شکل ۹ مقدار BSS به تفکیک اشیای مختلف مجموعه داده MSRC برای دو الگوریتم قطعه‌بندی SLIC با تعداد ۱۶۰ ناحیه و EGBS و اعمال روش پیشنهادی SDE به این دو الگوریتم آورده شده است.

دومین پارامتری که برای ارزیابی از آن استفاده شده است پارامتر Recall به ازای آستانه‌گذاری مقادیر مختلف BSS می‌باشد. این پارامتر عبارت است از شمارش تعداد اشیاء با در نظر گرفتن یک مقدار BSS که به عنوان یک حداقل مقدار آستانه تعریف می‌شود [۲].

$$Recall_T = \frac{Num\ of\ Class_i\ | \ BSS_i \geq T}{N} \quad i \in \{1, \dots, N\} \quad (3)$$

در رابطه‌ی ۳ متغیر T مقدار آستانه و N تعداد دسته شیء مجموعه داده را نشان می‌دهد. عبارت BSS_i نشان‌دهنده‌ی بیشترین مقدار همپوشانی برای دسته i ام از مجموعه داده است. در جدول ۱ مقادیر مختلف Recall بر اساس مقدار آستانه حداقل ۵۰٪ برای الگوریتم SLIC با تعداد ناحیه‌های مختلف، الگوریتم EGBS و اعمال الگوریتم SDE به هر دو الگوریتم قطعه‌بندی آورده شده است.

جدول ۱ مقادیر Recall بر اساس حداقل مقدار آستانه ۵۰٪ برای الگوریتم قطعه‌بندی و اعمال الگوریتم SDE به آن

N	Algorithm	Recall
۱۰	SLIC	۰٫۷۶
۱۰	SLIC-SDE	۰٫۹۰
۴۰	SLIC	۰٫۶۶
۴۰	SLIC-SDE	۱
۷۰	SLIC	۰٫۶۶
۷۰	SLIC-SDE	۰٫۹۵
۱۰۰	SLIC	۰٫۵۲
۱۰۰	SLIC-SDE	۰٫۹۵
۱۶۰	SLIC	۰٫۵۲
۱۶۰	SLIC-SDE	۰٫۸۵
-	EGBS	۰٫۹۵
-	EGBS-SDE	۱

¹ Area under the curve

با توجه به تعاریف ارائه شده در بالا می‌توان مقدار دقت (P_r) و فراخوانی (R_r) را برای هر تصویر محاسبه نمود و همچنین مقدار $F_{measure}$ را به دست آورد. این روابط در فرمول‌های ۸ تا ۱۰ آورده شده‌اند [۲۴].

$$P_r = \frac{|P_{11}|}{|P_{11}|+|P_{10}|} \quad (۸)$$

$$R_r = \frac{|P_{11}|}{|P_{11}|+|P_{01}|} \quad (۹)$$

$$F_{measure} = \frac{2 \times P_r \times R_r}{P_r + R_r} \quad (۱۰)$$

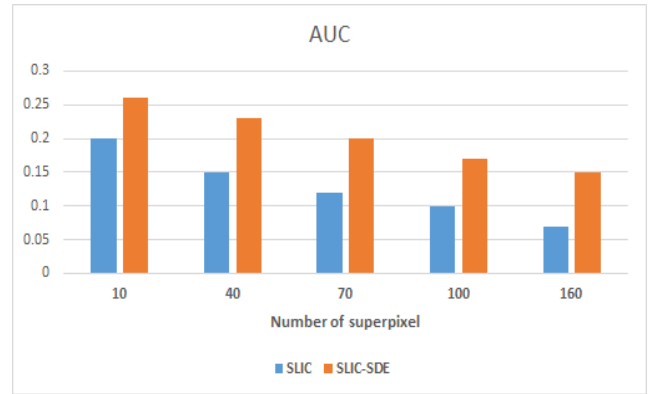
برای ارزیابی کامل‌تر، پارامتر^۱ RI برای بررسی کیفیت ناحیه‌ها محاسبه شده است که در ارزیابی کیفیت طبقه‌بندی معمولاً به کار گرفته می‌شود. چگونگی محاسبه‌ی آن در رابطه‌ی ۱۱ آورده شده است [۲۶].

$$RI = \frac{|P_{11}|+|P_{00}|}{|P|} \quad (۱۱)$$

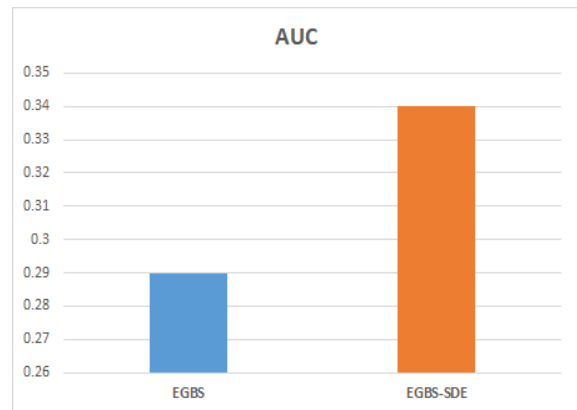
در جدول ۲ چهار پارامتر معرفی شده در بالا برای یک تصویر نمونه محاسبه شده است. بیشترین مقدار RI برای الگوریتم EGBS و SLIC (با تعداد ۱۰ ناحیه) به دست آمده است. با توجه به داده‌های به دست آمده مقدار $F_{measure}$ با اعمال الگوریتم SDE به الگوریتم قطعه‌بندی SLIC و EGBS به ترتیب به میزان ۱۶% و ۲۶% افزایش می‌یابد.

جدول ۲ نتایج دقت، فراخوانی، F-measure، RI و RI برای یک تصویر نمونه

	N	Precision	Recall	$F_{measure}$	RI
SLIC	۱۰	۰/۷۷	۰/۱۴	۰/۲۴	۰/۳۶
SLIC-SDE	۱۰	۰/۷۰	۱	۰/۸۲	۰/۷۰
SLIC	۴۰	۰/۸۶	۰/۰۳	۰/۰۶	۰/۳۱
SLIC-SDE	۴۰	۰/۵۶	۰/۱۲	۰/۲۰	۰/۳۱
SLIC	۷۰	۰/۸۸	۰/۰۲	۰/۰۴	۰/۳۱
SLIC-SDE	۷۰	۰/۹۳	۰/۰۴	۰/۰۸	۰/۳۲
SLIC	۱۰۰	۰/۸۸	۰/۰۱	۰/۰۲	۰/۳۰
SLIC-SDE	۱۰۰	۰/۹۲	۰/۰۲	۰/۰۴	۰/۳۰
SLIC	۱۶۰	۰/۸۹	۰/۰۰۸	۰/۰۱	۰/۳۰
SLIC-SDE	۱۶۰	۰/۹۳	۰/۰۱۷	۰/۰۳۳	۰/۳۰
EGBS	۵۲	۰/۸۵	۰/۴۱	۰/۵۶	۰/۵۳
EGBS-SDE	۵۲	۰/۷۳	۰/۹۳	۰/۸۲	۰/۷۱



شکل ۱۱ مقدار AUC محاسبه شده برای نمودار Recall بر اساس همپوشانی با حداقل مقدار ۵۰% و N تعداد سوپرپیکسل را نشان می‌دهد.



شکل ۱۲ مقدار AUC محاسبه شده برای نمودار Recall بر اساس همپوشانی با حداقل مقدار ۵۰% برای قطعه بندی EGBS

یکی دیگر از معیارهای ارزیابی مبتنی بر ناحیه‌بندی، دیدگاه طبقه‌بندی بر اساس همبستگی موجود در جفت پیکسل‌های تصویر است. در این دیدگاه، ناحیه‌بندی یک تصویر به مفهوم طبقه‌بندی همه‌ی جفت پیکسل‌ها به دو ناحیه است و اینکه آیا این جفت پیکسل به یک ناحیه تعلق دارد یا به ناحیه‌های مجزا تعلق دارد. اگر فرض کنید مجموعه‌ی تمام پیکسل‌های یک تصویر به صورت مجموعه‌ی $I = \{p_1, p_2, \dots, p_n\}$ باشد و مجموعه‌ی همه‌ی جفت پیکسل‌ها را با $\mathcal{P} = \{(p_i, p_j) \in I \times I \mid i < j\}$ فرض کنیم، آنگاه با داشتن دو ناحیه‌بندی S و S' مجموعه‌ی \mathcal{P} به چهار زیرمجموعه تقسیم می‌شود. یک جفت پیکسل (p_i, p_j) از مجموعه‌ی پیکسل‌ها در یکی از چهار زیرمجموعه زیر (روابط ۴ تا ۷) قرار می‌گیرد [۲۳]:

$$P_{11}: \text{قرار گرفتن جفت پیکسل در هر دو ناحیه } S \text{ و } S' \quad (۴)$$

$$P_{10}: \text{قرار گرفتن جفت پیکسل فقط در ناحیه } S \quad (۵)$$

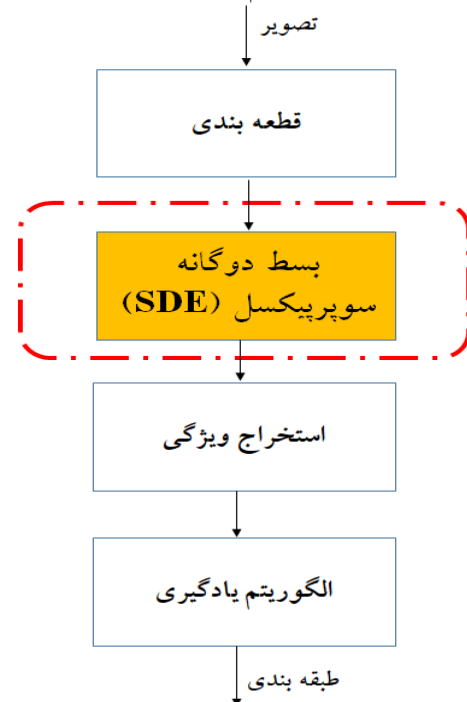
$$P_{01}: \text{قرار گرفتن جفت پیکسل فقط در ناحیه } S' \quad (۶)$$

$$P_{00}: \text{نبودن جفت پیکسل در ناحیه‌های } S \text{ و } S' \quad (۷)$$

¹ Rand index

۲-۴ طبقه‌بندی تصویر

یکی دیگر از مسائل مهم کاربردی در بینایی ماشین طبقه‌بندی تصویر است. برای حل مساله طبقه‌بندی تصویر در اینجا فرض بر این است که هر تصویر فقط دارای چندبرچسب است. هر برچسب نشان‌دهنده وجود شیء در تصویر می‌باشد و اطلاعاتی از مکان اشیاء در تصویر در دسترس نیست. همچنین ویژگی‌های اشیاء نیز مشخص نیست. در حقیقت مساله با فرض نظارت ضعیف^۱ بررسی می‌شود. از این رو، مساله‌ی طبقه‌بندی به یک یادگیری چندنمونه‌ای تبدیل می‌شود. معماری پیشنهادی برای طبقه‌بندی تصویر در شکل ۱۳ نشان داده شده است. بعد از استخراج ناحیه‌های تصویر با الگوریتم قطعه‌بندی در مرحله اول، الگوریتم بسط دوگانه سوپرپیکسل (SDE) به آن اعمال می‌شود. سپس در مرحله سوم ویژگی‌های سطح پایین برای تمام نواحی تصویر (که هر ناحیه معادل یک نمونه در الگوریتم یادگیری چندنمونه‌ای می‌باشد) استخراج می‌شوند. در مرحله پایانی با بکارگیری طبقه‌بند مناسب، طبقه‌بندی تصاویر انجام می‌شود.



شکل ۱۳ ساختار پیشنهادی برای طبقه‌بندی تصویر مبتنی بر الگوریتم پیشنهادی SDE که این واحد پیشنهادی در مرحله دوم افزوده شده است

بررسی هیستوگرام رنگ و همچنین هیستوگرام الگوی باینری محلی^۲ [۲۷] ناحیه محاسبه شده است. در مرحله پایانی یک طبقه‌بند مناسب برای آن اعمال می‌شود که در این مرحله الگوریتم یادگیری چندنمونه‌ای استفاده شده است. از میان الگوریتم‌های مختلف پیشنهاد شده مبتنی بر یادگیری چندنمونه‌ای، ما در این مقاله از MIL-Boost که توسط ویولا و همکاران [۲۸] معرفی شده است، استفاده می‌کنیم. این روش از توابع هزینه یادگیری چند نمونه‌ای که با چارچوب AnyBoost ترکیب شده است، استفاده می‌کند.

به‌طور تصادفی مجموعه‌ی داده به دو گروه مجزا برای مرحله-ی آموزش و آزمایش تقسیم شده‌اند. همه‌ی تصاویر در ابتدا به اندازه‌ی یکسانی تغییر اندازه داده شده، سپس الگوریتم SLIC برای استخراج سوپرپیکسل به هر تصویر اعمال شده است. برای ارزیابی الگوریتم تعداد ناحیه‌ها در دو حالت مختلف ۷۰ و ۱۰ بررسی شده‌اند. برای هر ناحیه یک بردار ویژگی ۸۰ عضوی، محاسبه می‌شود. سپس الگوریتم یادگیری چندنمونه‌ای بر روی داده‌های آموزش و آزمایش اعمال شده و مقدار دقت اندازه‌گیری شده است. برای بررسی بهتر تاثیر بسط ناحیه، طبقه‌بندی تصویر در دو حالت بدون بسط و بسط با اندازه‌ها و حالت‌های متفاوت ارزیابی شده است. در شکل ۱۴ مقدار دقت شناسایی برای حالت «fixed-mode» با دو مقدار متفاوت P (مقدار پیکسل) و با استفاده از الگوریتم یادگیری MIL-Boost بر اساس گروه‌های مختلف مجموعه داده MSRC ترسیم شده است. نتایج به‌دست آمده از «all-mode» و به‌کارگیری الگوریتم MIL-Boost در شکل ۱۵ نشان داده شده است. همچنین الگوریتم SDE برای تعداد ۱۰ ناحیه در حالت all-mode بررسی شده که در شکل ۱۶ نتیجه‌ی آن مشاهده می‌شود. هر سه شکل ۱۴، ۱۵ و ۱۶ تاثیرگذاری استفاده از الگوریتم SDE را تایید می‌کند.

۵ بحث

باتوجه به اهمیت ناحیه‌بندی ادراکی در حل مسائل شناسایی دیداری، بهبود الگوریتم‌های قطعه‌بندی می‌تواند تاثیر به‌سزایی در اینگونه مسائل داشته باشد. الگوریتم پیشنهادی SDE با توجه به ویژگی‌های الگوریتم‌های قطعه‌بندی قادر به بهبود میزان همپوشانی اشیای تصاویر است. دلیل آن را می‌توان تفاوت اجزای مختلف اشیاء در پارامترهایی مانند رنگ و بافت دانست. در نتیجه با بسط دادن سوپرپیکسل به ناحیه‌های همسایه این ایراد به میزان مناسبی برطرف می‌شود. علاوه بر این در الگوریتم قطعه‌بندی تعداد ناحیه‌های استخراج شده از تصویر در مقایسه با تعداد اشیاء و مفاهیم تصویر معمولاً خیلی بیشتر است. این نقصان موجب افزایش میزان حافظه‌ی مصرفی و کاهش سرعت محاسبات می‌گردد. بنابراین اجرای بسط سوپرپیکسل می‌تواند یک راه حل عملی برای این

در مرحله اولیه، تصویر با تعدادی ناحیه‌ی غیرهمپوشان نمایش داده می‌شود. برای استخراج این نواحی از الگوریتم SLIC استفاده می‌کنیم. حال بعد از به‌دست آمدن تمام ناحیه‌های تصویر، هر ناحیه به‌طور مجزا بر اساس یکی از مجموعه شرایط پیشنهادی بسط داده می‌شود. برای استخراج ویژگی روش‌های مختلفی وجود دارد و می‌توان از توصیف‌گرهای مختلف مانند هیستوگرام رنگ، گرایان لبه، گشتاور، بافت و دیگر موارد استفاده کرد. در این

² Local binary pattern

¹ Weakly supervised

جزئی و به صورت خطی را نتیجه می‌دهد که قابل چشم‌پوشی می‌باشد.

جدول ۳ مقدار دقت طبقه‌بندی برای چند دسته شیء از مجموعه داده MSRC برای SLIC و SLIC-SDE با ۷۰ سوپرپیکسل

Object	SLIC	SLIC-SDE		
		(p=4)	(p=8)	All-mode
Aeroplane	۰,۸۵	۰,۸۸	۰,۹۳	۰,۸۴
Bicycle	۰,۸۸	۰,۹۴	۰,۹۴	۰,۹۵
Bird	۰,۷۲	۰,۸	۰,۷۹	۰,۸۵
Boat	۰,۸۱	۰,۸۳	۰,۸۰	۰,۸۷
Body	۰,۴۹	۰,۵۱	۰,۶۰	۰,۵۲
Cow	۰,۵۶	۰,۸۵	۰,۸۴	۰,۸۲
Dog	۰,۷۴	۰,۹۰	۰,۸۶	۰,۸۲
Grass	۰,۶۹	۰,۷۵	۰,۷۵	۰,۷۶
Road	۰,۶۴	۰,۶۸	۰,۷۱	۰,۶۸
Water	۰,۶۸	۰,۷۶	۰,۷۶	۰,۷۸

جدول ۴ مقدار دقت طبقه‌بندی برای چند دسته شیء از مجموعه داده MSRC برای SLIC و SLIC-SDE با ۱۰ سوپرپیکسل

Object	SLIC	SLIC-SDE
		All-mode
Sheep	۰,۴۵	۰,۹۰
Bicycle	۰,۵۳	۰,۹۱
Face	۰,۶۶	۰,۷۷
Boat	۰,۷۶	۰,۸۸
Body	۰,۵۹	۰,۶۸
Cow	۰,۸۲	۰,۸۷
Dog	۰,۵۴	۰,۸۶
Grass	۰,۵۶	۰,۷۷
Road	۰,۵۴	۰,۶۹
Water	۰,۷۳	۰,۸۱

۶ نتیجه‌گیری

با توجه به اهمیت و تاثیرگذاری روش نامزدهای شیء در حل مسائل بینایی ماشین، بهبود مساله‌ی قطعه‌بندی تصویر در مرحله‌ی ابتدایی محاسبات نامزد شیء اهمیت به‌سزایی در کارایی آن خواهد داشت. در این مقاله یک الگوریتم بسط دوگانه سوپرپیکسل SDE به منظور بهبود کارایی قطعه‌بندی مبتنی بر ناحیه پیشنهاد و بررسی شد. این الگوریتم شامل دو حالت مختلف با هدف احاطه‌ی اشیا و مختلف تصویر است. این الگوریتم می‌تواند در الگوریتم‌های دیگر قطعه‌بندی استفاده شود که در این مقاله دو الگوریتم قطعه‌بندی EGBS و SLIC مورد ارزیابی قرار گرفته‌اند. در مجموع الگوریتم SDE پارامترهای مختلف ارزیابی مرتبط با قطعه‌بندی را بهبود بخشید. همچنین این الگوریتم برای حل مساله‌ی طبقه‌بندی تصویر بر اساس الگوریتم یادگیری چندنمونه‌ای MIL مورد مطالعه قرار گرفت که بهبود نتایج آن نیز به وضوح قابل مشاهده است.

مشکل باشد. در حقیقت بسط یک ناحیه به ناحیه‌های مجاور بر اساس قوانین خاص و با توجه به شرایط گفته شده می‌تواند در استخراج ناحیه‌ی بزرگتر و دقیق‌تر که متناسب با اشیا و مفاهیم یک تصویر هستند، موثرتر باشد.

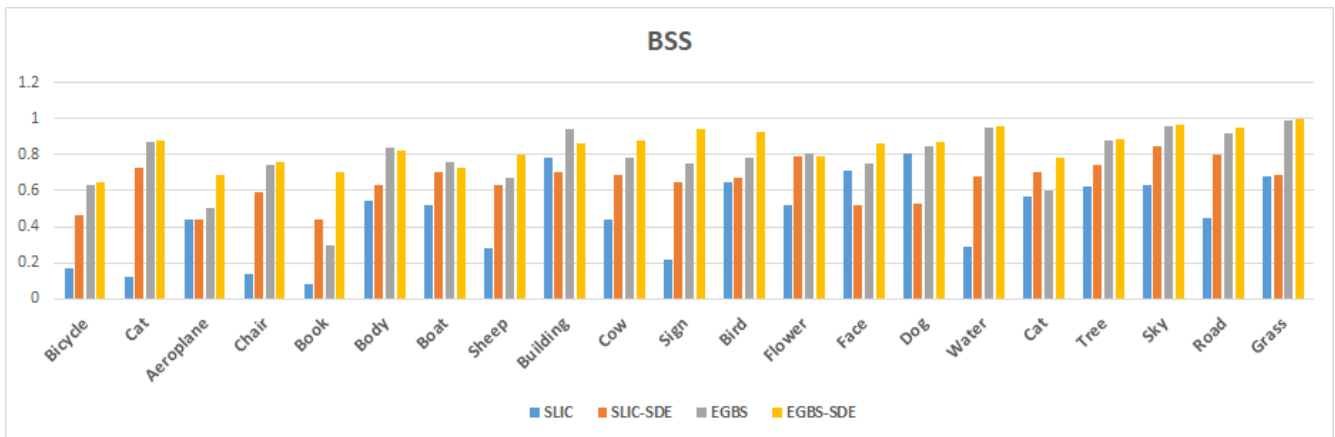
همانگونه که نتایج ارزیابی در بخش قبلی بیان شد، به کارگیری الگوریتم پیشنهادی SDE افزایش همپوشانی را برای اشیا و مختلف موجود در مجموعه داده MSRC نتیجه می‌دهد. همچنین، میانگین پارامتر RI با استفاده از روش پیشنهادی در الگوریتم قطعه‌بندی SLIC به مقدار ۷٪ بهبود می‌یابد و این پارامتر برای الگوریتم EGBS به مقدار ۱۸٪ افزایش داشته است و برای دیگر پارامترها مانند Recall، دقت و فراخوانی و $F_{measure}$ افزایش کارایی برای هر دو الگوریتم SLIC و EGBS نشان داده شده است.

همچنین با به کارگیری الگوریتم SDE در حل مساله طبقه‌بندی تصاویر افزایش کارایی قابل مشاهده است. البته ثابت ماندن میزان حافظه‌ی مورد نیاز برای ذخیره‌سازی بردارهای ویژگی تصویر بعد از بسط ناحیه‌های سوپرپیکسل از امتیازات دیگر این روش است. برای مثال در شرایطی که در الگوریتم SLIC تعداد ۱۰ سوپرپیکسل در نظر گرفته شود در مجموع مقدار بهبود دقت طبقه‌بندی ۱۱٪ به دست می‌آید و اگر تعداد سوپرپیکسل ۷۰ فرض شود مقدار افزایش دقت برابر با ۴,۲٪ است. روش SDE برای اشیا و مفاهیم ارائه شده در جدول ۳ افزایش قابل توجهی دارد. مقدار این افزایش برای ۱۰ تا از اشیا این مجموعه داده ۱۱٪ می‌باشد. در این جدول تعداد سوپرپیکسل‌های هر تصویر ۷۰ فرض شده است. در جدول ۴ مقدار بهبود دقت را برای تعدادی از اشیا با توجه به در نظر گرفتن ۱۰ سوپرپیکسل در مرحله‌ی ابتدایی قطعه‌بندی نشان می‌دهد. در مجموع مقدار بهبود به‌طور متوسط برابر با ۲۰٪ شده است. در این مقاله هدف اصلی بررسی تاثیر روش پیشنهادی برای ناحیه‌بندی است که به خوبی پارامترهای مهم و شناخته شده در ارزیابی قطعه‌بندی و همین‌طور طبقه‌بندی تصویر را بهبود می‌دهد. نکته‌ی دیگری که در این الگوریتم وجود دارد اینکه بهبود این نتایج پیچیدگی ناچیزی را به سیستم تحمیل می‌کند. برای الگوریتم‌های EGBS و SLIC استفاده شده در این مقاله پیچیدگی زمانی به ترتیب عبارت است از $O(N \log N)$ و $O(N)$ که مقدار N تعداد پیکسل‌های تصویر را بیان می‌کند و مقدار پیچیدگی مرحله‌ی دوم الگوریتم نیز به صورت مقدار خطی می‌باشد که به تعداد ناحیه‌های استخراج شده در مرحله‌ی اول وابسته است. اگر تعداد این ناحیه‌ها M در نظر گرفته شود با به کارگیری الگوریتم EGBS در مرحله‌ی اول و با توجه به اینکه $M \ll N \log N$ است مقدار پیچیدگی برابر با $O(N \log N + M) \cong O(N \log N)$ خواهد بود. اگر در مرحله‌ی ابتدایی از الگوریتم قطعه‌بندی SLIC استفاده شود با توجه به اینکه $M \ll N$ است مقدار پیچیدگی آن برابر با $O(N + M) \cong N$ خواهد شد. در مجموع الگوریتم پیشنهادی با بهبود پارامترهایی از قطعه‌بندی افزایش پیچیدگی زمانی خیلی

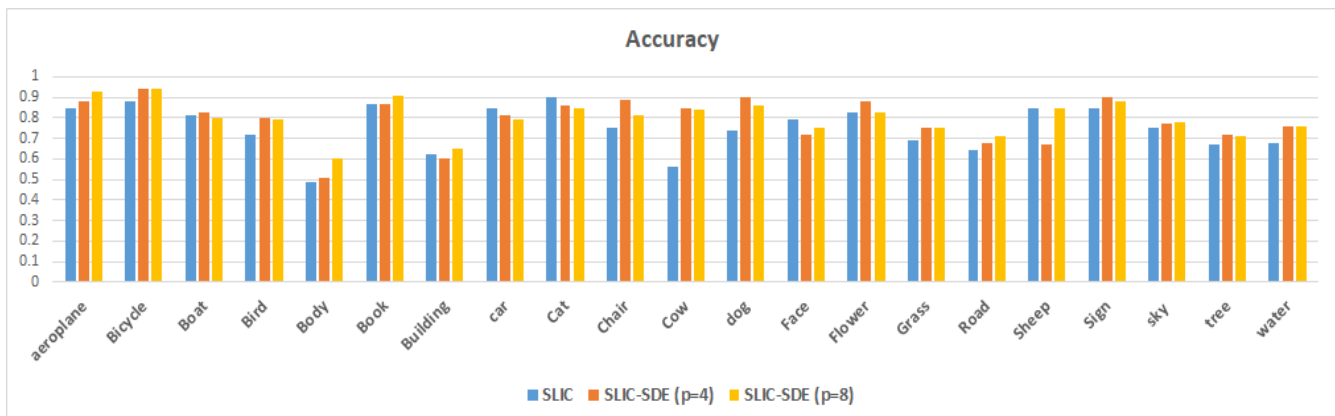
۷ سپاسگزاری

سپاسگزاریم. همچنین از مشورت‌های آقای دکتر علی جان نشاری از دانشگاه برکلی قدردانی می‌گردد.

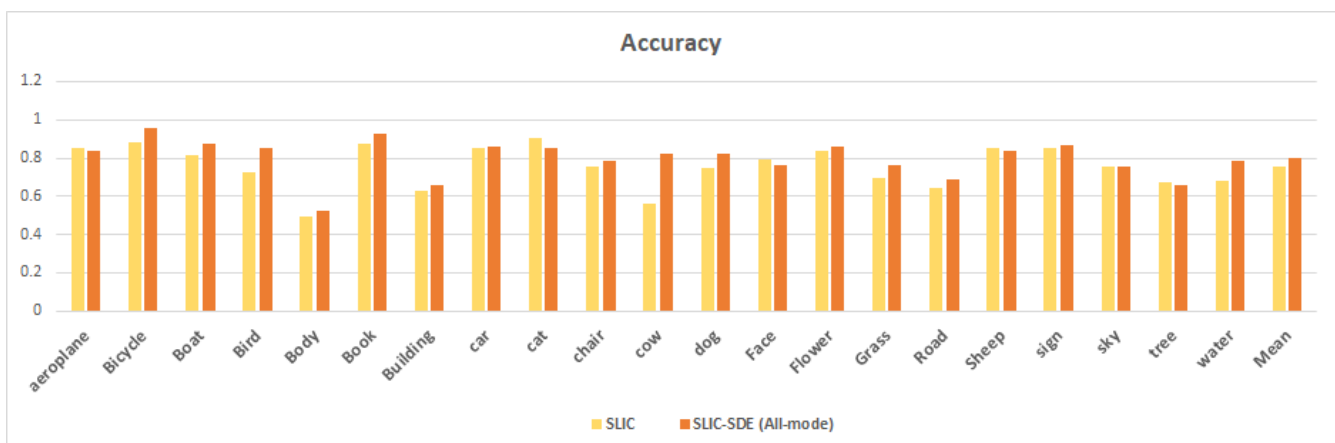
از مرکز تحقیقات دانشکده مهندسی برق و کامپیوتر دانشگاه برکلی برای در اختیار گذاشتن جزئیات الگوریتم یادگیری چندنمونه ای



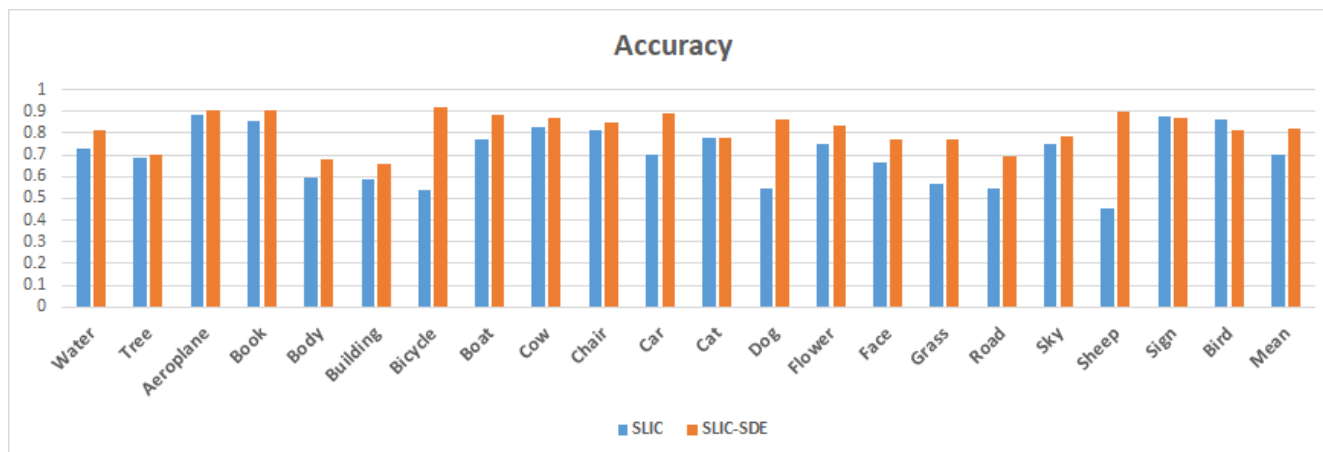
شکل ۹ مقدار BSS در دسته اشیاء مختلف مجموعه داده MSRC برای الگوریتم‌های SLIC (با ۱۶۰ ناحیه) و EGBS و اعمال SDE به آن‌ها



شکل ۱۴ مقادیر دقت برای دسته اشیاء مختلف در SLIC-SDE و SLIC با ۷۰ ناحیه در حالت fixed-mode با دو مقدار $p=4, p=8$



شکل ۱۵ مقادیر دقت طبقه‌بندی به تفکیک دسته اشیاء مختلف در SLIC-SDE و SLIC با ۷۰ ناحیه در حالت all-mode



شکل ۱۶ مقادیر دقت طبقه‌بندی به تفکیک دسته اشیاء مختلف در SLIC و SLIC-SDE با فرض ۱۰ ناحیه در حالت all-mode

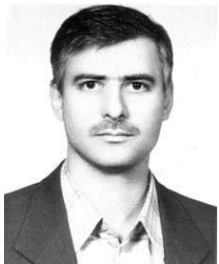
- [10] G.T. Papadopoulos, V. Mezaris, S. Dasiopoulou, and I. Kompatsiaris, "Semantic image analysis using a learning approach and spatial context," In International conference on semantic and digital media technologies, 2006.
- [11] J. Carreira, and C. Sminchisescu, "CPMC: Automatic Object Segmentation Using Constrained Parametric Min-Cuts", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34, pp. 1312-1328, 2012.
- [12] J. Dai, K. He, and J. Sun, "Convolutional Feature Masking for Joint Object and Stuff Segmentation," In IEEE conference on computer vision and pattern recognition, 2015.
- [13] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," In IEEE conference on computer vision and pattern recognition, 2014.
- [14] B. Hariharan, P. Arbeláez, R. Girshick, J. Malik. "Simultaneous Detection and Segmentation," In European conference on computer vision, 2014.
- [15] Wei. Xia, Csaba Domokos, Jian Dong, Loong-Fah Cheong and Shuicheng Yan, "Semantic Segmentation without Annotating Segments," In IEEE conference on computer vision, 2013.
- [16] H. Zhu, F. Meng, J. Cai, and S. Lu, "Beyond Pixels: A Comprehensive Survey from Bottom-Up to Semantic Image Segmentation and Cosegmentation," Journal of visual communication and image representation, vol. 34, pp. 12-27, 2016.
- [17] J. Shi, and J. Malik, "Normalized Cuts and Image Segmentation," IEEE transactions on pattern analysis and machine intelligence, vol. 22, pp. 888-905, 2000.
- [18] A. Vedaldi, and S. Soatto, "Quick shift and kernel methods for mode seeking," In European conference on computer vision, 2008.
- [19] D. Comaniciu, and P. Meer, "Mean shift: A robust approach toward feature space analysis," IEEE Transactions on Pattern analysis and machine intelligence, vol. 24, pp. 603-619, 2002.

مراجع

- [1] C. Gu, J. J. Lim, P. Arbeláez, and J. Malik, "Recognition using regions," IEEE Conference on computer vision and pattern recognition, pp. 1030-1037, 2009.
- [2] I. Endres, and D. Hoiem, "Category independent object proposals," Computer vision- ECCV, pp. 575-588, 2010.
- [3] P. Arbeláez, B. Hariharan, C. Gu, S. Gupta, L. Bourdev, and J. Malik, "Semantic Segmentation using regions and parts," In IEEE conference on computer vision and pattern recognition, 2012.
- [4] N. Dalal, and B. Triggs, "Histograms of Oriented Gradients for Human Detection," In IEEE conference on computer vision and pattern recognition, 2005.
- [5] P. F. Felzenszwalb, R. Girshick, D. McAllester and D. Ramanan, "Object Detection with Discriminatively Trained Part Based Models," IEEE Transaction pattern analysis and machine intelligence, vol. 32, pp. 1627-1645, 2010.
- [6] J. Pont-Tuset, P. Arbelaez, J. T. Barron, F. Marques, and J. Malik, "Multiscale Combinatorial grouping for image segmentation and object proposal generation," IEEE transactions on pattern analysis and machine intelligence, vol. 39, pp. 128-140, 2017.
- [7] J. Uijlings, K. van de Sande, T. Gevers, and A. Smeulders, "Selective search for object recognition," International journal of computer vision, vol. 104, pp. 154-171, 2013.
- [8] B. Alexe, T. Deselaers, V. Ferrari, "What is an Object?," In IEEE conference on computer vision and pattern recognition, 2010.
- [9] M. H. Quinn, E. Conser, J. M. Witte, and M. Mitchell, "Semantic Image Retrieval via Active Grounding of Visual Situations", In IEEE conference on computer vision and pattern recognition, 2017.



مریم تقی‌زاده مدرک کارشناسی و کارشناسی ارشد خود را در رشته مهندسی کامپیوتر گرایش سخت‌افزار به ترتیب از دانشگاه‌های اصفهان و تهران اخذ نموده‌اند. ایشان هم‌اکنون در مقطع دکتری تخصصی در رشته مهندسی کامپیوتر در دانشگاه رازی کرمانشاه مشغول به تحصیل می‌باشند. زمینه‌های تحقیقاتی وی بینایی ماشین، یادگیری عمیق، شناسایی اشیاء و تشخیص الگو است.



عبدالله چاله‌چاله مدرک کارشناسی (مهندسی برق سخت‌افزار) و کارشناسی ارشد (مهندسی کامپیوتر-نرم‌افزار) خود را از دانشگاه صنعتی شریف ایران و دکتری مهندسی کامپیوتر خود را از دانشگاه ولونگونگ استرالیا دریافت نمود. ایشان هم‌اکنون به عنوان عضو هیأت علمی در دانشگاه رازی کرمانشاه مشغول به کار است. زمینه‌های علمی مورد علاقه ایشان پردازش تصویر و ویدئو و نیز روش‌های جدید تعامل انسان و کامپیوتر می‌باشد.

- [20] P. F. Felzenszwalb, and D. P. Huttenlocher, "Efficient Graph-based Image Segmentation," International journal of computer vision, vol. 59, pp. 167-181, 2004
- [21] R. Achanta, A. Shaji, K. Smith A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC Superpixels compared to state-of-the-art superpixel methods," IEEE Transactions on pattern analysis and machine intelligence, vol. 34, pp.2274-2282, 2012.
- [22] J. Shotton, J. Winn, C. Rother, and A. Criminisi, "TextonBoost: Joint Appearance, Shape and Context Modeling for Multi-class Object Recognition and Segmentation," In European conference on computer vision, 2006.
- [23] J. Pont-Tuset and F. Marques, "Measures and Meta-Measures for the Supervised Evaluation of Image Segmentation," In IEEE conference on computer vision and pattern recognition, 2013.
- [24] D. R. Matrin, *An Empirical approach to grouping and segmentation*, Ph.D. thesis, EECS Department, University of California, Berkeley, 2003.
- [25] R. Szeliski, *Computer vision: Algorithms and applications*, Springer, 2010.
- [26] W. Rand. "Objective criteria for the evaluation of clustering methods," Journal of the American Statistical Association, vol. 66, pp. 846-850, 1971.
- [27] T. Ojala, M. Pietikäinen, and D. Harwood, "Performance Evaluation of Texture measures with classification Based on Kullback Discrimination of Distribution", Pattern Recognition, 1994.
- [28] P. Vial, J. C. Platt, and C. Zhang, "Multiple Instance Boosting for Object detection," Advances in neural information processing systems, pp. 1417-1424, 2006.