

دریافت مقاله: ۱۳۹۵/۰۴/۲۰

پذیرش مقاله: ۱۳۹۶/۰۱/۲۷

## انتخاب ناحیه‌های کاندیدا در سیستم‌های تشخیص و شناسایی اشیاء

علی قنبری سرخی<sup>۱\*</sup>، حمید حسن‌پور<sup>۲</sup>، منصور فاتح<sup>۳</sup>

<sup>۱\*</sup> دانشجوی دکتری، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شاهرود، شاهرود، ایران

ali.ghanbari289@gmail.com

<sup>۲</sup> استاد، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شاهرود، شاهرود، ایران

h.hassanpour@shahroodut.ac.ir

<sup>۳</sup> استادیار، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شاهرود، شاهرود، ایران

mansoor\_fateh@yahoo.com

### چکیده

در تحقیقات انجام شده در سال‌های اخیر، به دست آوردن ناحیه‌های کاندیدا به عنوان یک مرحله اساسی و مهم در سیستم‌های تشخیص و شناسایی اشیاء موجود در تصویر معرفی شده است. به دست آوردن این ناحیه‌ها مانند یک تنگنا بوده و بیشترین بار محاسباتی را در این نوع از سیستم‌ها دارد. در این خصوص، انتخاب روش مناسب و سریع می‌تواند در بهبود عملکرد سیستم‌های تشخیص بسیار حائز اهمیت باشد. در این مقاله، به مرور کارهای انجام شده در این زمینه پرداخته شده و چندین روش مشهور و محبوب مورد استفاده در سیستم‌های شناسایی قدرتمند معرفی شده است. همچنین به مقایسه و ارزیابی روش‌های مطرح بر روی مجموعه داده‌های استاندارد PASCAL VOC، ImageNet، COCO و COCO پرداخته شده است. در بین روش‌های مورد ارزیابی، روش ناحیه کاندیدا گروه‌بندی ترکیبی بر پایه چندین مقیاس (MCG) با الگوریتم شناسایی شبکه عصبی کانولوشن بر پایه ناحیه (R-CNN)، بهترین نتایج را داشته است. این روش عملکردی در حدود ۵۷ درصد، ۵۴ درصد و ۴۱ درصد بر روی مجموعه داده‌های PASCAL VOC 2007، ImageNet 2013 و COCO 2014 نشان داده است.

واژه‌های کلیدی: ناحیه کاندیدا، R-CNN سریع، ابرپیکسل، شیء بودن، تشخیص و شناسایی اشیاء.

## ۱. مقدمه

نزدیک به زمان بلادرنگ در استفاده از شبکه‌های عمیق در صورتی که زمان محاسبه ناحیه‌های پیشنهادی را در نظر نگیریم، به دست آورده است [۵]. در واقع در این شبکه، زمان مربوط به شناسایی ناچیز است.

روش‌های نواحی کانیدها به طور معمول بر ویژگی‌های ارزان و سریع تکیه می‌کنند. جستجوی انتخابی [۶] یکی از معروف‌ترین روش‌های پیشنهاد ناحیه است که به صورت حریرصانه، ابرپیکسل‌هایی<sup>۸</sup> را که بر اساس ویژگی‌های سطح پایین (مانند میزان شدت روشنایی و رنگ در فضاها رنگی متفاوت) طراحی شده، ادغام می‌کند. اما این روش در مقایسه با شبکه‌های شناسایی کارآمد مانند R-CNN سریع [۴] کندتر است. روش EdgeBoxes پیشنهادشده در [۷]، اخیراً یک سازش<sup>۹</sup> بین کیفیت کانیدها و زمان را فراهم کرده است. باوجود این، مراحل تولید نواحی کانیدها، بار محاسباتی زیادی (از لحاظ زمان اجرا) در شبکه‌های تشخیص ایجاد می‌کنند. باید دقت کرد که معمولاً CNN مبتنی بر ناحیه بر اساس GPU توسعه یافته شده است، در حالی که روش‌های ناحیه کانیدها بر روی CPU پیاده‌سازی می‌شوند. در واقع بسیاری از روش‌های کانیدها قابلیت موازی‌سازی را به خوبی ندارند؛ به همین دلیل چنین مقایسه‌ای بین روش‌های پیشنهاد ناحیه و شناسایی شیء در زمان اجرا نابرابر است. همان‌طور که اشاره شد، بهبود عملکرد سیستم‌های تشخیص و شناسایی اشیاء وابستگی زیادی به روش‌های ناحیه کانیدها دارند. در این خصوص، در این مقاله به کارهای انجام‌شده در این زمینه می‌پردازیم و روش‌های موجود را معرفی و مقایسه می‌کنیم. از طرفی بیان شد که R-CNN برای شناسایی اشیاء به‌عنوان یک سیستم مشهور و محبوب بوده و در بسیاری از کارهای صنعتی و تحقیقاتی، مورد استقبال محققان قرار گرفته است. به همین دلیل در این مقاله، برای شناسایی اشیاء، از این روش به‌همراه روش‌های نواحی کانیدها استفاده شده است.

سیستم‌های تشخیص اشیاء که در سال‌های اخیر پیشنهاد شده‌اند، به الگوریتم‌های ناحیه‌های کانیدها<sup>۱</sup> برای تخمین مکان اشیاء بسیار وابسته‌اند. از مهم‌ترین شبکه‌های تشخیص می‌توان به SPPnet [۱] و R-CNN<sup>۳</sup> سریع [۲] اشاره کرد که در سال‌های اخیر، مورد توجه بسیاری از محققان و کاربرهای صنعتی قرار گرفته است. در این روش‌ها نمی‌توان سراسر تصویر را به دلیل محدودیت‌های زمانی و حالت‌های مختلف از اشیاء جستجو کرد. در این خصوص، ابتدا باید ناحیه‌های کانیدها وجود شیء استخراج شوند. پیشرفت‌های انجام‌شده سبب بهبود زمان اجرای شبکه‌های تشخیص شده و همچنین نشان داده شده تعیین نواحی کانیدها مهم‌ترین تنگنای<sup>۴</sup> محاسباتی این گونه روش‌هاست. نمونه‌ای از ناحیه‌های کانیدها برای بررسی وجود شیء مورد نظر در شکل (۱) نشان داده شده است. همان‌طور که در شکل نشان داده شده، سیستم تشخیص شیء، انتخاب نهایی خود را از ناحیه‌های کانیدها انجام می‌دهد.

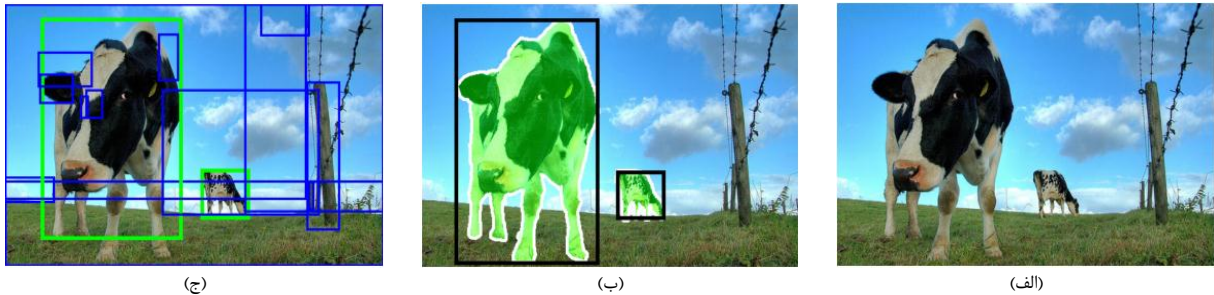
در مرجع [۲]، یک شبکه پیشنهاد ناحیه کانیدها<sup>۵</sup> (RPN) معرفی شده است. در این شبکه، ویژگی‌های با کانولوشن‌های مختلف استخراج‌شده از سراسر تصویر به اشتراک گذاشته می‌شود. در نتیجه، زمان مربوط به پیشنهادهای ناحیه را قابل چشم‌پوشی می‌سازد. RPN به طور همزمان مرزهای اشیاء و امتیاز<sup>۶</sup> شیء بودن<sup>۷</sup> را در هر موقعیت از تصویر پیش‌بینی می‌کند. اگرچه در مرجع [۳] محاسبات بالا در شبکه‌های عصبی کانولوشن بر پایه ناحیه به‌عنوان یک اصل بیان شده، این هزینه‌ها با اشتراک کانیدها به دست‌آمده از کانولوشن، به شدت در حال کاهش است [۱ و ۴]. در آزمایش‌های انجام‌شده در سال‌های اخیر، R-CNN سریع [۴]، نرخ

## 8. Superpixels

ابرپیکسل به نواحی اتمیک به هم پیوسته و دارای مفهوم واحد در تصویر گفته می‌شود.

## 9. tradeoff

1. Region proposal
2. Spatial pyramid pooling network
3. Region-convolutional neural network
4. bottleneck
5. Region Proposal Network
6. Score
7. Objectness



شکل (۱): نمونه‌ای از ناحیه کاندیدا: الف. تصویر اصلی، ب. اشیاء موجود در تصویر، ج. ناحیه‌های کاندیدا [۸]

به‌طور مستقیم از کانتورهای لبه<sup>۳</sup> (EC) [۱۱, ۱۲] تقسیم می‌شوند. از مهم‌ترین روش‌های موجود در این حوزه می‌توان به موارد زیر اشاره نمود:

- جستجوی انتخابی<sup>۴</sup> [۶ و ۸]: در این روش از ادغام ابرپیکسل‌ها، برای تولید کاندیداها استفاده می‌شود. این روش پارامترهای یادگیری ندارد و از ویژگی‌ها و تابع‌های شباهت برای ادغام ابرپیکسل‌ها استفاده می‌کند. در سال‌های اخیر این روش، به صورت گسترده به عنوان روش کاندیدا برای بسیاری از آشکارسازهای شیء مورد استفاده قرار گرفته شده است.

در مقاله<sup>۵</sup> [۶] جستجوی انتخابی با ترکیب استراتژی جستجوی جامع<sup>۵</sup> و قطعه‌بندی ارائه شده است. در این روش مانند روش‌های قطعه‌بندی، در نمونه‌گیری از ساختار تصویر استفاده شده است. همچنین از جستجوی جامع برای محدود کردن تمامی موقعیت‌های ممکن از اشیاء استفاده شده است. نتایج به دست آمده در این روش، روی مجموعه کوچکی با وابستگی به داده و مستقل از کلاس، بیشترین دقت در مکان‌یابی حاصل شده است. در واقع، در مجموعه داده‌هایی با کلاس‌های از پیش مشخص نبوده، این روش عملکرد خوبی داشته است. همچنین در این روش، آموزشی برای کلاس‌ها به صورت جداگانه صورت نگرفته است. آزمایش‌های انجام شده دقتی برابر با ۹۹ درصد و متوسط میانگین بهترین

در بخش دوم مقاله به معرفی روش‌های پیشنهاد ناحیه‌های کاندیدا پرداخت می‌شود. در قسمت سوم مقاله اشاره می‌شود که کنترل تعداد نواحی کاندیدا می‌تواند در مقایسه بین روش‌ها، بسیار مهم باشد. در این باره، در این قسمت به تحلیل آن پرداخته شده است. معرفی مجموعه داده استفاده شده و مقایسه بین روش‌های موجود، در قسمت‌های چهارم و پنجم این پژوهش بیان شده است.

## ۲. روش‌های پیشنهاد ناحیه کاندیدا

همان‌طور که در قسمت قبل اشاره شد، استخراج ناحیه‌های کاندیدا تأثیر بسیار زیادی در دقت و سرعت روش‌های تشخیص و شناسایی اشیاء در تصاویر دارند. به‌طور کلی، می‌توان روش‌های پیشنهاد نواحی کاندیدا را به چهار دسته روش‌های مبتنی بر گروه‌بندی، امتیازدهی به پنجره کاندیدا، جایگزین و مرجع تقسیم کرد که در ادامه به تشریح این روش‌ها پرداخته شده است.

### ۱.۲. روش‌های مبتنی بر گروه‌بندی

روش‌های مبتنی بر گروه‌بندی برای استخراج ناحیه‌های کاندیدا در سه گروه براساس چگونگی تولید کاندیدا، دسته‌بندی می‌شوند. به‌طور کلی، روش‌های تولید کاندیدا بر پایه گروه‌بندی به گروه‌بندی ابرپیکسل (SP) [۹]، حل کردن مسائل برش گراف چندگانه<sup>۱</sup> (GC) با seeds<sup>۲</sup> [۱۰] متنوع یا

منظور از Seeds، ابرپیکسل‌های استخراج شده از طریق نمونه‌گیری انرژی‌محور است. در مقاله<sup>۱۰</sup> [۱۰] یک تابع انرژی بر پایه شباهت رنگ بین مرزها و هیستوگرام رنگ ابرپیکسل‌ها معرفی شده است.

3. Edge contours

4. <http://koen.me/research/downloads/SelectiveSearchCodeIJCV.zip>

5. Exhaustive

### 1. Multiple graph cut

منظور از برش در گراف، تقسیم رئوس گراف به دو زیرمجموعه ناتهی جدا از هم S و V/S است.

### 2. Seeds: Superpixels extracted via energy-driven sampling

ویژگی‌ها شامل، رنگ در فضای RGB، شدت روشنایی<sup>۱</sup> (I) در تصاویر، کانال رنگی RG به‌همراه RGB، HSV، I، RGB نرمال شده، H در HSV و C [۱۳] هستند.

شکل (۳) نمونه‌ای از نتایج به‌دست‌آمده از روش جستجوی انتخابی را نشان می‌دهد. در هر تصویر محدوده قرمز ناحیه کاندیدا و محدوده سبز مربوط به ناحیه هدف است. مقداری که زیر هر عکس ذکر شده، میزان همپوشانی این دو محدوده را نشان می‌دهد.



Bike: 0.863

Person: 0.882

شکل (۳): نمونه‌ای از خروجی الگوریتم [۶]. محدوده قرمز، ناحیه کاندیدا روش جستجوی انتخابی و محدوده سبز، ناحیه هدف است

• پریم تصادفی<sup>۲</sup>: در مرجع [۱۴] از شباهت ویژگی‌ها مانند روش جستجوی انتخابی استفاده شده است. در این روش یک فرایند ادغام جدید به‌صورت تصادفی برای ابرپیکسل‌ها معرفی شده است. در این فرایند ادغام همه احتمالات ممکن آموزش داده شده است. با این روش، سرعت به‌صورت قابل ملاحظه‌ای افزایش پیدا کرده است.

• جستجو محلی و سراسری<sup>۳</sup>: در مرجع [۱۵] یک استراتژی مشابه جستجوی انتخابی، برای ادغام ابرپیکسل‌ها پیشنهاد شده است. ولی از ویژگی‌های مختلف و بیشتری استفاده کرده است. در مرحله بعدی، قطعات<sup>۴</sup> تولیدشده به عنوان Seeds به‌کار گرفته شده و برای حل کردن برش

همپوشانی بین نواحی پیشنهادی و نواحی صحیح ۰/۸۷۹ در ۱۰۰۹۷ موقعیت را از خود نشان داده است. در این روش از یک الگوریتم گروه‌بندی سلسله مراتبی استفاده شده است. ناحیه در مقایسه با پیکسل اطلاعات بیشتری دارد؛ به همین دلیل، در این مقاله از ویژگی‌ها بر پایه ناحیه بیشتر استفاده شده است. در واقع مجموعه‌ای از پیکسل‌ها که ویژگی‌های یکسانی دارند استفاده می‌شوند. برای انتخاب مجموعه‌ای از ناحیه‌های هم‌پوشان با اشیاء از روش سریع [۹] و [۱۱] استفاده شده است. الگوریتم کلی در شکل (۲) آورده شده است.

#### Algorithm 1: Hierarchical Grouping Algorithm

**Input:** (colour) image

**Output:** Set of object location hypotheses  $L$

Obtain initial regions  $R = \{r_1, \dots, r_n\}$  using [9]

Initialise similarity set  $S = \emptyset$

**foreach** Neighbouring region pair  $(r_i, r_j)$  **do**

    Calculate similarity  $s(r_i, r_j)$

$S = S \cup s(r_i, r_j)$

**while**  $S \neq \emptyset$  **do**

    Get highest similarity  $s(r_i, r_j) = \max(S)$

    Merge corresponding regions  $r_i = r_i \cup r_j$

    Remove similarities regarding  $r_i: S = S \setminus s(r_i, r_*)$

    Remove similarities regarding  $r_j: S = S \setminus s(r_*, r_j)$

    Calculate similarity set  $S_i$  between  $r_i$  and its neighbours

$S = S \cup S_i$

$R = R \cup r_i$

Extract object location boxes  $L$  from all regions in  $R$

#### شکل (۲): الگوریتم گروه‌بندی سلسله‌مراتبی در جستجوی انتخابی [۶]

همان‌طور که اشاره شد، از روش ارائه‌شده در مرجع [۹] برای تولید ناحیه‌های اولیه استفاده شده است. سپس از یک الگوریتم حریم‌بندی ناحیه‌ها باهم استفاده می‌شود. در این الگوریتم، ابتدا شباهت بین همه ناحیه‌های همسایه محاسبه می‌شود و در گام بعدی، ناحیه‌هایی با بیشترین شباهت، باهم ترکیب می‌شوند. این کار تا زمانی ادامه می‌یابد که همه تصاویر به یک ناحیه واحد تبدیل شوند. در این مقاله اشاره شد که ناحیه‌های تولیدشده می‌توانند در صحنه‌ها و شرایط روشنایی متفاوت باشند. در همین راستا الگوریتم گروه‌بندی در فضاهای رنگی متفاوت با ویژگی‌های مستقل و متفاوت اجرا شده است. در این روش از معیار شباهت وابسته به چهار پارامتر رنگ، بافت، اندازه و میزان همپوشانی ناحیه‌ها نسبت به هم، برای گروه‌بندی ناحیه‌ها استفاده شده است. این

1. intensity

2. RandomizedPrim

<https://github.com/smanenfr/rp/archive/master.zip>

3. Rantalankila

<https://github.com/Cloud-CV/object-proposals/tree/master/rantalankilaSegments>

4. segments

توضیح دهد. در این مقاله، برجستگی سطح شیء توسط اطلاعات به دست آمده از ابرپیکسل‌ها با یک معیار اندازه‌گیری جدید محاسبه شده است.

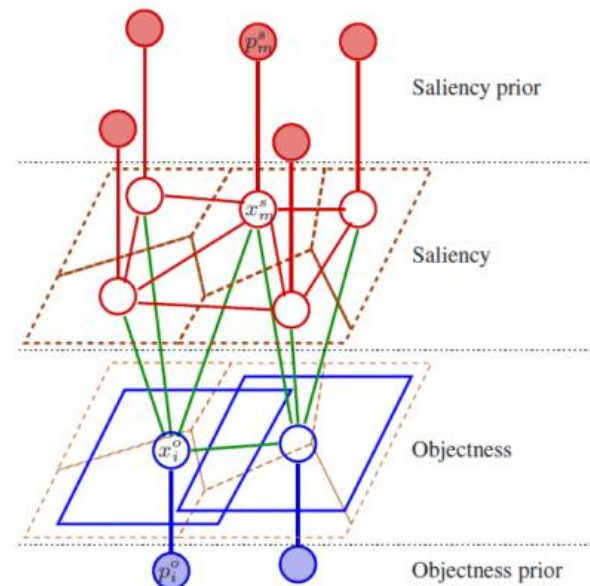
• قطعه‌بندی خودکار با برش‌های کمینه<sup>۴</sup> (CPMC) [۱۷]، [۱۸] از قطعه‌بندی اولیه اجتناب می‌کند و برش‌های گراف را با Seedهای مختلف مستقیماً روی پیکسل محاسبه می‌کند. نتایج قطعه‌بندی با حجم بزرگی از ویژگی‌ها رتبه‌بندی می‌شوند.

در مقاله [۱۷] یک روش جدید برای تولید و رتبه‌بندی اشیاء قابل قبول در تصویر با فرایند پایین به بالا<sup>۵</sup> و نشانه‌های سطح متوسط ارائه شده است. منظور از نشانه‌های سطح متوسط، میزان شباهت بافت، روشنایی، انرژی کانتور بین ناحیه‌ها و درون ناحیه‌هاست. اشیاء به صورت قطعه‌بندی از تصویر استخراج شده است. در این روش، قطعه‌بندی به صورت خودکار و بدون دانش اولیه از خواص کلاس اشیاء بوده است.

برای این منظور، از حل دنباله‌ای از محدوده پارامتری برش کمینه بر روی شبکه منظمی<sup>۶</sup> از تصویر استفاده شده است. در گام بعدی، برای پیش‌بینی میزان قابل قبول بودن قطعه انتخابی، یک مدل پیوسته برای رتبه‌بندی اشیاء آموزش داده می‌شود. این روش برای قطعه‌بندی سطح پایین در مجموعه داده VOC09 عملکرد خوبی را از خود نشان داده است. این روش متوسط بهترین همپوشانی ۰/۷۸ با ۱۵۷ قطعه را دارد. در CMPC برای هر تصویر، مجموعه‌ای از پیکسل‌ها متعلق به پیش‌زمینه فرض می‌شوند. این پیش‌زمینه به عنوان Seeds معرفی می‌شود. سپس برای هر مجموعه، چندین سطوح از پیش‌زمینه استخراج می‌شود. هزینه‌های متفاوتی به همه پیکسل‌های باقی‌مانده اختصاص داده می‌شود. اما بعضی از پیکسل‌های در امتداد مرز تصویر، جز Seed منفی می‌شوند. از الگوریتم‌ها بر پایه گراف به منظور قطعه‌بندی این Seedها استفاده شده است. همچنین میزان شباهت بین پیکسل‌های همسایه به عنوان وزن گراف در

گراف از CPMC<sup>۱</sup> (در قسمت بعد توضیح داده می‌شود) برای تولید کاندیداهای بیشتر استفاده شده است.

• چانگ [۱۶] ترکیبی از برجستگی<sup>۲</sup> و شیء‌بودن با مدل گرافیکی، برای ادغام ابرپیکسل در قطعه‌بندی تصویر/ پس‌زمینه معرفی کرده است. مراحل کلی از مدل گرافی پیشنهادی در شکل (۴) نشان داده شده است.



شکل (۴): مراحل کلی الگوریتم چانگ [۱۶] شامل نمونه‌گیری تصافی جدید و انتصاب مقدار برجستگی به هر پنجره، تخمین شیء‌بودن است.

این روش با نمونه‌گیری تصادفی از پنجره‌های زیاد شروع می‌شود. در گام نخست این روش، تعداد زیادی پنجره، به صورت تصادفی با اندازه‌های متغیر انتخاب می‌شود. به هر پنجره مقدار شیء‌بودن و برای هر پیکسل یا ابرپیکسل مقدار برجستگی اختصاص داده می‌شود. برای ارتباط شیء‌بودن و برجستگی، مقداری به عنوان سطح برجستگی شیء<sup>۳</sup> برای هر پنجره معرفی شده است. این مقدار برای نمایش برجستگی جسم زمینه در هر پنجره استفاده می‌شود. از شیء‌بودن برای تخمین برجستگی استفاده شده است. همچنین از برجستگی برای تخمین شیء‌بودن استفاده شده است. مقدار شیء‌بودن هر پنجره در صورتی بالا خواهد بود که برجستگی سطح شیء به خوبی بتواند بسیاری از مقادیر برجستگی پیکسل‌های را

4. Cpmc: Automatic object segmentation using constrained parametric min-cuts  
[http://www.maths.lth.se/matematiklth/personal/sminchis/code/cpmc/cpmc\\_release1.gz](http://www.maths.lth.se/matematiklth/personal/sminchis/code/cpmc/cpmc_release1.gz)

5. Down-Top

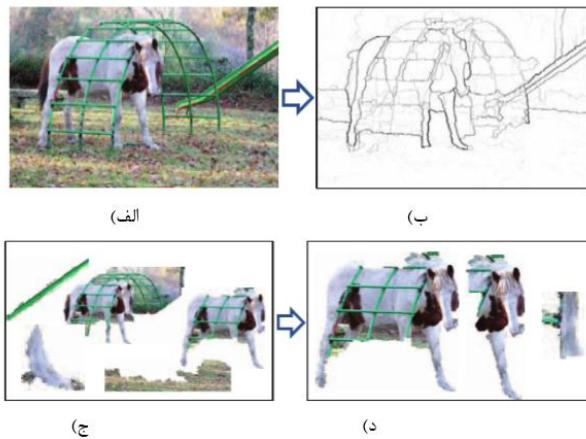
6. cues

7. grid

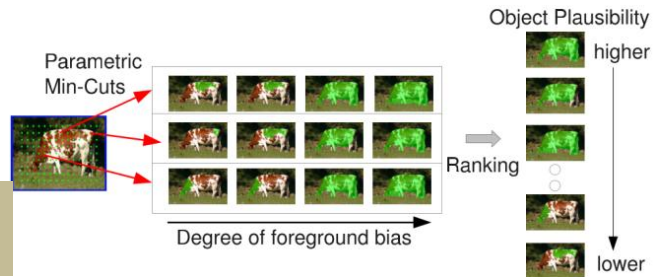
1. Constrained Parametric Min-Cuts

2. Saliency

3. Object-level saliency



شکل (۶): مراحل مختلف از روش [۱۹]، الف. تصویر ورودی، ب. قطعه‌بندی سلسله‌مراتبی، ج. ناحیه‌های کانیدها، د. رتبه‌بندی ناحیه‌ها



شکل (۵): روال پیشنهادشده در [۱۷]

• استحکام داده‌ها<sup>۴</sup> [۲۳] یک روش بهبودیافته از CPMC است که سرعت محاسبه Seedها را با محاسبه مجدد در سراسر مسائل برش گراف چندگانه، به صورت قابل ملاحظه‌ای کاهش می‌دهد. همچنین از روش آشکارسازی سریع لبه استفاده شده است [۱۲ و ۲۴]. این الگوریتم بر روی مجموعه داده PASCAL VOC 2012 مورد ارزیابی قرار گرفته شده است. نتایج آزمایش‌های گزارش شده در این مقاله نشان‌دهنده بهبود روش پیشنهادی نسبت به روش‌های دیگر است. بار محاسباتی این روش، بر روی ماشین با مشخصات 3.2GHz Intel i7-3930K با ۸ هسته در حدود ۲-۴ ثانیه در CPU برای هر تصویر است.

• روش Geodesic<sup>۵</sup> [۲۲] توسط یک روش قطعه‌بندی تصویر پیشنهادی در مرجع [۱۲] شروع می‌شود. طبقه‌بندی‌ها از مکان Seeds برای تبدیل فاصله Geodesic<sup>۶</sup> استفاده می‌کنند. در این مقاله، Seeds توسط طبقه‌بندی بهینه برای کشف اشیاء، آموزش داده شده است.

• در مراجع [۱۹ و ۲۰] یک قطعه‌بندی سلسله‌مراتبی از مرزهای انسداد<sup>۲</sup> و حل برش گراف با Seedها و پارامترهای متفاوت، برای تولید قطعات استفاده شده است. کانیدها براساس طیف گسترده‌ای از نشانه‌ها رتبه‌بندی شده است. این عمل در واقع سبب تشویق تنوع انتخابات می‌شود.

مراحل کلی پیشنهادشده در مقاله [۱۹] در شکل (۶) نشان داده شده است. در این روش، تصویر ورودی براساس روش سلسله‌مراتبی ارائه شده در [۲۱] قطعه‌بندی می‌شوند. این قطعه تولیدشده، ناحیه‌های Seed تصویر را به وجود می‌آوردند. در گام بعدی، با استفاده از این Seedها و ویژگی‌های مانند مرز، رنگ و بافت، مجموعه متنوعی از نواحی تولید می‌شود. این نواحی می‌توانند به عنوان قطعات شیء معرفی شوند. در نهایت با روش یادگیری ساخت یافته، ناحیه‌های به دست آمده رتبه‌بندی می‌شوند که به احتمال زیاد، مناطق با امتیاز بالاتر مربوط به اشیاء مختلف هستند. نتایج به دست آمده بر روی مجموعه داده BSDS<sup>۳</sup> و PASCAL VOC 2011 نشان‌دهنده توانایی این روش برای پیدا کردن اشیاء بیشتر با نواحی با اندازه کوچک در پنجره‌های کانیدهاست.

#### 4. Rigor

[http://cpl.cc.gatech.edu/projects/RIGOR/resources/rigor\\_src.zip](http://cpl.cc.gatech.edu/projects/RIGOR/resources/rigor_src.zip)

۵. منظور از فاصله Geodesic، اندازه کوتاه‌ترین مسیر بین نودهاست.

[http://www.philkr.net/papers/2014-10-01-eccv/gop\\_1.3.zip](http://www.philkr.net/papers/2014-10-01-eccv/gop_1.3.zip)

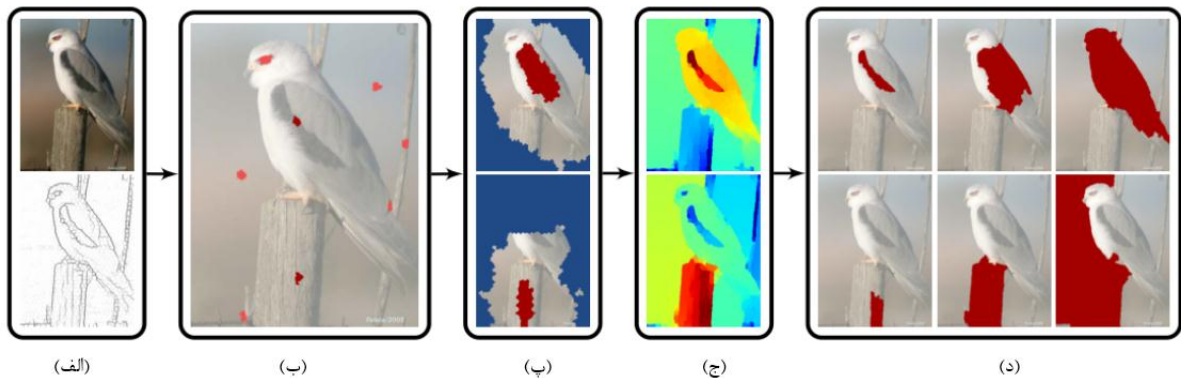
۶. منظور از تبدیل فاصله Geodesic، معیاری برای اندازه‌گیری کوتاه‌ترین مسیر بین مجموعه‌ای از نودها تا هر نود

#### 1. Endres

<https://github.com/Cloud-CV/object-proposals/tree/master/endres/proposals>

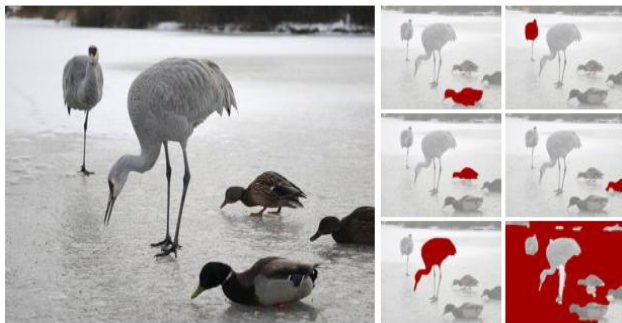
#### 2. occlusion

#### 3. Berkeley Segmentation Data Set



شکل (۷): مراحل کلی از روش Geodesic مطرح شده در مرجع [۲۲] الف. تصویر ورودی و بیش‌قطعه‌بندی آن به‌همراه نقشهٔ احتمال مرز، ب. مکان Seeds، پ. ماسک پیش‌زمینه و پس‌زمینهٔ تولیدشده برای دو Seeds، ج. استفاده از تبدیل فاصله Geodesic (SGDB) برای این ماسک‌ها، د. کاندیداهای پیشنهادی با توجه به محاسبه مجموعه‌های سطوح بحرانی در هر SGDB

یک تابع خطی دیگر برای پس‌زمینه آموزش داده شده است. نمونه‌ای از نتیجهٔ به‌دست‌آمده از روش پیشنهادی، بر روی مجموعه‌دادهٔ PASCAL VOC 2012 در شکل (۸) نشان داده شده است.



شکل (۸): اشیاء کاندیدا (نواحی قرمز) توسط الگوریتم [۲۲]

آزمایش‌های انجام‌شده در این مقاله، عملکرد بالا از لحاظ دقت و سرعت را روی مجموعه‌دادهٔ PASCAL VOC 2012 گزارش داده است.

• گروه‌بندی ترکیبی بر پایهٔ چندین مقیاس<sup>۶</sup> (MCG) [۲۸]: یک الگوریتم سریع برای محاسبهٔ قطعه‌بندی سلسله‌مراتبی چندین-مقیاسه<sup>۷</sup> که در مرجع [۱۲] معرفی شده، ارائه کرده است. قطعات براساس استحکام لبه<sup>۸</sup> ادغام می‌شوند. در نتیجه اشیاء براساس نشانه‌هایی مانند اندازه، مکان، شکل و استحکام لبه رتبه‌بندی شده‌اند.

ایدهٔ اصلی در این مقاله، معرفی مجموعه‌های سطوح بحرانی<sup>۱</sup> در تبدیل فاصلهٔ Geodesic است. تبدیل فاصلهٔ Geodesic برای مکان‌های Seeds در تصویر محاسبه شده است. تبدیل فاصله، در زمان نزدیک برخط<sup>۲</sup> محاسبه می‌شود. در این روش، محاسبهٔ اندازه‌گیری کوتاه‌ترین مسیر بین مجموعه‌ای از نودها تا هر نود ناچیز است. همچنین محاسبهٔ هر تبدیل برای تولید کاندیدها در مقیاس‌های متفاوت استفاده شده است. از این‌رو، استفاده از خط‌لوله<sup>۳</sup> در محاسبات، می‌تواند بسیار مؤثر باشد.

همان‌طور که در شکل (۷) نشان داده شده، در این روش ابرپیکسل‌های تصاویر ورودی با قطعه‌بندی محاسبه می‌شوند و نقشهٔ احتمال مرز<sup>۴</sup> به‌عنوان وزن گراف در نظر گرفته می‌شود. در واقع در این گراف هر نود به یک ابرپیکسل مربوط بوده و هر یال اتصال‌های بین ابرپیکسل‌های مجاور است. وزن یال‌ها میزان این تبدیل، به صورت کلی پس‌زمینه و پیش‌زمینه را مشخص می‌کند [۲۷]. هر مجموعه سطح<sup>۵</sup> از این تکنیک می‌تواند یک قطعه از تصویر را محدود کند که می‌تواند به‌عنوان اشیاء کاندیدا در نظر گرفته شوند. منظور از مجموعه سطح، مرز اشیاء موجود در پیش‌زمینه است. در این مقاله پیش‌زمینه و پس‌زمینهٔ به دست‌آمده از روش قبل با یک روش بر پایهٔ یادگیری بهبود داده شده است. در روش پیشنهادی، یک تابع خطی برای پیش‌زمینه و

6. Multiscale combinatorial grouping  
<https://github.com/jponttuset/mcg/archive/v2.0.zip>  
 7. Multi-scale  
 8. Edge Strength

1. Critical level sets  
 2. online  
 3. Pipeline  
 4. Boundary probability map  
 5. Level set

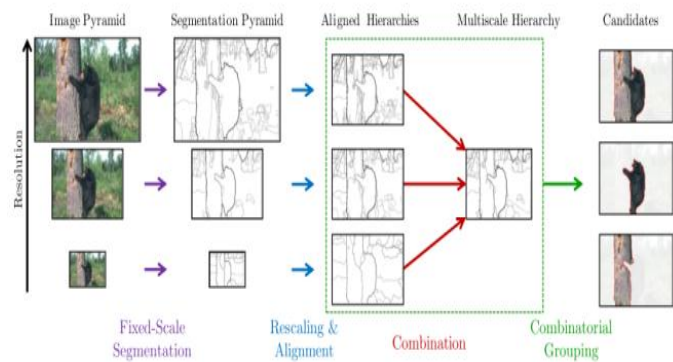


شکل (۱۰): نمونه‌ای از خروجی مرجع [۲۸]. قسمت بالای تصویر:

به ترتیب از سمت چپ: تصویر اصلی، نواحی هدف، قطعه‌بندی

سلسله‌مراتبی در چندین مقیاس، قسمت پایین تصویر: بهترین اشیاء

کاندیدا



شکل (۹): مراحل کلی از روش پیشنهادی در [۲۸]

تولید می‌کنند. در ادامه به معرفی بعضی از این روش‌ها می‌پردازیم.

- شی‌بودن [۲۹ و ۳۰] یکی از اولین و بهترین روش‌های شناخته‌شده در بحث روش‌های کاندیداهاست. یک مجموعه اولیه از کاندیداها از محل‌های برجسته در تصویر انتخاب می‌شود. این کاندیداها براساس چندین نشانه مانند رنگ، لبه‌ها، محل، اندازه و استحکام "Superpixel straddling (SS)" امتیازدهی می‌شوند. برای بررسی استحکام SS معیاری معرفی شده که میزان اشتراک بین ابرپیکسل و پنجره پیشنهادی را محاسبه می‌کند.

- Rahtu [۳۱] از مجموعه‌ای از ناحیه‌های کاندیدا با انتخاب تصادفی استفاده می‌کند. این نواحی توسط ابرپیکسل‌های منحصربه‌فرد، جفت یا سه‌تایی به دست می‌آیند. این ابرپیکسل‌ها چندین محدوده با انتخاب تصادفی دارند. استراتژی امتیازدهی بازبینی شده و بهبودهایی در آن ارائه شده است.

- Bing [۳۲] از یک طبقه‌بند خطی ساده با ویژگی‌های لبه استفاده می‌کند. در این روش، از پنجره لغزان استفاده شده است. سرعت تخمین در این روش بسیار بالاست. در این روش، طبقه‌بند تأثیر اندکی دارد و می‌تواند عملکردی مشابه با

مراحل کلی روش پیشنهادی در شکل (۹) نشان داده شده است. این روش، ابتدا از هرمی در چندین وضوح<sup>۱</sup> تصویر شروع می‌شود. در گام بعدی، قطعه‌بندی سلسله‌مراتبی در هر مقیاس، به صورت مستقل صورت می‌گیرد. سپس تصاویر حاصل از سلسله‌مراتبی چندگانه تراز می‌شوند و تصاویر تولیدی در یک تصویر قطعه‌بندی شده با چندین مقیاس تنها ترکیب می‌شوند. در نهایت با گروه‌بندی اجزای حاصل، فهرست امتیازاتی از اشیاء کاندیدا با جستجو در ناحیه‌های فضای ترکیبی محاسبه می‌شود.

این روش بر روی مجموعه داده PASCAL VOC 2012 و BSDS500 مورد ارزیابی قرار گرفته است و عملکرد مناسبی از این روش گزارش شده است. نمونه‌ای از نتایج به دست آمده در شکل (۱۰) نشان داده شده است.

## ۲.۲. روش‌های امتیاز دادن به پنجره کاندیدا

یک روش جایگزین برای تولید کاندیداها، امتیازدهی هر پنجره کاندیداهاست. در واقع هر تصویر پنجره‌گذاری شده و در گام بعدی به هر پنجره امتیازدهی براساس احتمال شی‌بودن پنجره کاندیدا صورت می‌گیرد. در مقایسه با روش‌های گروه‌بندی، این روش‌ها معمولاً فقط مرز محدوده‌ها را برگشت می‌دهند و سریع‌تر هستند. در مواردی که نمونه‌گیری پنجره با تعداد بالا باشد، این روش‌ها معمولاً پیشنهادهای با دقت محلی‌سازی پایین

2. <http://groups.inf.ed.ac.uk/calvin/objectness/objectness-release-v۲,۲.zip>

3. [http://www.ee.oulu.fi/research/imag/object\\_detection/ObjectnessICCV\\_ver۰۲.zip](http://www.ee.oulu.fi/research/imag/object_detection/ObjectnessICCV_ver۰۲.zip)

4. <https://github.com/tfzhou/BINGObjectness/archive/master.zip>



PASCAL VOC 2007 تخمین زده می‌شوند سپس ۵/۰ درصد از کوچک‌ترین و بزرگ‌ترین مقادیر نادیده گرفته می‌شوند و توزیع تخمین‌زده شده ۹۹ درصد از داده‌ها را پوشش می‌دهد.

• روش گاوسی<sup>۸</sup>: به همین ترتیب، یک توزیع گوسین چندمتغیره برای موقعیت مرکز محدوده مرزبندی، ناحیه ریشه مربع و لگاریتم نسبت ابعاد تخمین‌زده می‌شود. در گام بعدی، مقدار میانگین و واریانس در مجموعه آموزشی محاسبه می‌شود. در نهایت کاندیداها از این توزیع‌ها نمونه‌برداری شده‌اند.

• روش پنجره لغزان: در این روش پنجره در یک شبکه (گرید) منظم مکان‌یابی می‌شود که برای آشکارسازی شیء در پنجره لغزان رایج است. تعداد درخواست کاندیداها براساس اندازه پنجره توزیع شده است. برای هر اندازه پنجره به صورت یکنواخت پنجره‌ها مکان‌یابی شده‌اند. این روش از پیاده‌سازی مراجع [۳۲ و ۳۳] الهام گرفته شده است.

• روش ابریکیسل: همان‌طور که در قسمت‌های قبل بیان شد، ابریکیسل تأثیر مهمی روی رفتار روش‌های پیشنهادی دارند. باتوجه‌به تعریف مرجع [۳۸]، ابریکیسل به نواحی اتمیک<sup>۹</sup> متراکم یا به‌هم‌پیوسته<sup>۱۰</sup> و ادراکی<sup>۱۱</sup> معنادار<sup>۱۲</sup> در گفته می‌شود.

### ۳. کنترل تعداد کاندیداهای پیشنهادی

برای مقایسه جامع بین روش‌های مطرح‌شده باید تعداد کاندیداهای ارائه‌شده در هر تصویر توسط روش‌هایی کنترل شوند؛ زیرا برای مقایسه صحیح بین روش‌های موجود باید تعداد کاندیداهای ارائه‌شده یکسان باشد. در بسیاری از موارد، روش‌ها با ارائه تعداد کاندیداها بین صد تا صدهزار مورد ارزیابی قرار می‌گیرند. همچنین برخی از روش‌ها امتیاز کاندیداها را ارائه داده و در مقابل بعضی دیگر این توانایی را

زمانی را که به تصویر نگاه نمی‌شود [۳۳] داشته باشد. این روش‌های مستقل از تصویر، با نام CrackingBing معروف هستند.

• EdgeBoxes [۷] همچنین از یک الگو پنجره لغزان بزرگ<sup>۲</sup> شروع می‌شود. اما براساس تخمین مرز اشیاء که توسط ساختار جنگل تصمیم [۱۴ و ۳۴] به‌دست آمده، ساخته شده است و یک گام پالایش<sup>۳</sup> بعدی را برای افزایش محلی‌سازی اضافه می‌کند. در این روش، پارامتری یاد گرفته نمی‌شود.

### ۲.۳. روش‌های کاندیدا جایگزین<sup>۴</sup>

• ShapeSharing [۳۵] یک روش غیر پارامتریک<sup>۵</sup> و مبتنی‌بر داده است که شکل‌های اشیاء را از نمونه‌های درون تصاویر آزمون توسط انطباق لبه‌ها استخراج می‌کند. در نتیجه نواحی در مرحله بعد ادغام می‌شوند و به‌وسیله حل کردن برش گراف اصلاح می‌شوند.

• Multibox [۳۶ و ۳۷] یک شبکه عصبی را برای برگرداندن تعداد ثابتی از کاندیداها به صورت مستقیم آموزش داده است. در واقع شبکه بر روی تصویر ورودی لغزانه نشده است. نویسنده بهترین نتایج بر روی دادگان ImageNet را گزارش داده است.

### ۲.۴. روش‌های پایه یا مرجع<sup>۶</sup> در پیشنهاد کاندیدا

بعضی از روش‌ها به‌عنوان روش پایه و نقاط مرجع معرفی می‌شوند. تمامی این روش‌ها، در قسمت‌های قبل مورد استفاده قرار گرفته‌اند. از روش‌های زیر، به‌عنوان روش پایه استفاده می‌شود.

• روش یکنواخت<sup>۷</sup>: در این روش برای تولید کاندیدا، به‌صورت یکنواخت موقعیت مرکز محدوده مرزبندی، ناحیه ریشه مربع و لگاریتم نسبت ابعاد، نمونه‌برداری می‌شوند. در گام بعدی، بازه‌ای از این پارامترها با مجموعه آموزش

1. <https://github.com/pdollar/edges/archive/master.zip>

2. coarse

3. refinement

4. Alternative Proposal Methods

5. Non-parametric

6. Baseline

7. Uniform

8. Gaussian

9. atomic

10. compact

11. perceptually

12. meaningful

بعضی از روش‌ها کاندیداهای تکراری تولید می‌کنند که آن‌ها را باید از مجموعه انتخابی در هر روش پاک نمود. مقایسه بین روش‌های ارائه شده از لحاظ کنترل پذیر بودن تعداد کاندیداها در جدول (۱) نشان داده شده است. همچنین در این جدول مقایسه‌ای بین روش‌ها از لحاظ نوع روش و خروجی حاصل صورت گرفته است.

#### ۴. مجموعه داده

در سال‌های اخیر، تحقیقات و آزمایش‌های بسیاری در زمینه انتخاب ناحیه‌های کاندیدا، تشخیص و شناسایی اشیاء بر روی تصاویر انجام شده است. معمولاً این تحقیقات بر روی مجموعه داده استاندارد بوده که ارزیابی روش‌های پیشنهادی را با دقت بیشتری امکان‌پذیر می‌سازد. از مهم‌ترین مجموعه داده‌های استفاده شده در سال‌های اخیر می‌توان به PASCAL، ImageNet و MS COCO اشاره کرد. در ادامه به معرفی هریک از مجموعه داده‌ها پرداخته شده و ویژگی هریک به تفصیل بیان می‌شود.

ImageNet: این مجموعه داده بزرگ‌ترین و گسترده‌ترین مجموعه تصاویر تهیه شده از وب است که برای تشخیص و شناسایی اشیاء مورد استفاده قرار گرفته شده است. این مجموعه شامل تقریباً ۳/۲ میلیون تصویر در ۵۲۴۷ کلاس متفاوت است. در تحقیقاتی که تاکنون انجام شده، تعداد محدودی از این مجموعه داده، مورد استفاده قرار گرفته است. مجموعه داده‌ای که در این تحقیق مورد بررسی قرار می‌گیرد، مجموعه اعتبارسنجی ImageNet 2013 است که شامل ۲۰۰ کلاس متفاوت و تقریباً بالغ بر ۲۰۰۰۰۰ تصویر از این کلاس‌ها است.

PASCAL: این مجموعه داده شامل ۲۰ کلاس متفاوت است که تقریباً شامل ۵۰۰۰ تصویر نامحدود در حالت‌های مختلف از این کلاس‌هاست.

MS COCO: دو مجموعه داده معرفی شده (ImageNet و PASCAL) برای مقایسه اگرچه از نظر تعداد کلاس‌ها

ندارند. البته نمی‌توان در همه روش‌ها به صورت صریح و روشن تعداد کاندیداها را کنترل کرد. در روش‌هایی که امتیاز کاندیدا را ارائه می‌دهند،  $k$  برترین ناحیه کاندیدا انتخاب می‌شوند. ولی در برخی از روش‌ها، نمی‌توان کنترل مستقیم بر روی تعداد کاندیداها داشت و همچنین امتیاز کاندیداها نیز مشخص نمی‌شوند. در این روش‌ها می‌توان با کنترل غیرمستقیم به  $k$  کاندیدا دسترسی داشت. در این خصوص نیاز است که پارامترهای دیگر را تغییر داد تا به این کنترل دسترسی پیدا کرد. به همین سبب تعداد کاندیداهای تولید شده با پارامترهای متفاوت در زیرمجموعه‌ای از تصویر ذخیره می‌شود. در نهایت با استفاده از یک درونیابی خطی بین تنظیم‌های پارامترها،  $k$  کاندیدا انتخاب می‌شوند. در بعضی از روش‌ها که هیچ کنترلی روی تعداد کاندیداها امکان‌پذیر نیست،  $k$  نمونه به صورت تصادفی انتخاب می‌شود.

جدول (۱): مقایسه‌ای بین روش‌های معرفی شده. علامت +/- به معنی، دارا بودن یا نبودن خاصیت مشخص است و علامت \* در کنترل تعداد کاندیدا، نشان‌دهنده کنترل غیرمستقیم است.

روش	نوع روش	خروجی قطعات	خروجی امتیاز	کنترل
Bing[18]	پنجره لغزان	-	+	+
CPMC[19]	گروه‌بندی	+	+	*
EdgeBox[20]	پنجره لغزان	-	+	+
Endres[21]	گروه‌بندی	+	+	+
Geodesic[22]	گروه‌بندی	+	-	*
MCG[23]	گروه‌بندی	+	+	*
Objectness[24]	پنجره لغزان	-	+	+
Rahtu[25]	پنجره لغزان	-	+	+
RandPrim[26]	گروه‌بندی	+	-	+
Rantalankila[27]	گروه‌بندی	+	-	*
Rigor[29]	گروه‌بندی	+	-	*
SelectiveSearch[30]	گروه‌بندی	+	+	*
Gussian	-	-	-	+
SlidingWindow	-	-	-	+
Superpixels	-	+	-	-
Uniform	-	-	-	+

۶۰ درصد از داده‌ها برای آموزش و ۴۰ درصد باقی‌مانده برای آزمون در نظر گرفته شده‌اند. در واقع نتایج ارائه‌شده بر پایه نتایج حاصل از مجموعه‌داده آزمون است. در روش‌های مطرح‌شده، روش‌های Bing, MCG, Rahtu, EdgeBoxes, Randomized Selective Search, Endres, CPMC و Prims نیاز به آموزش داشته و مرحله آموزش به صورت جداگانه انجام شده است. شکل (۱۱) نتایج به دست آمده از روش R-CNN سریع با روش‌های پیشنهاد کاندیدای متفاوت در هر کلاس را نشان می‌دهد. باتوجه به نتایج به دست آمده روش‌های EdgeBox و MCG در این مجموعه داده بر روی اکثر کلاس‌ها دارای بهترین نتایج بوده و عملکرد مناسبی را از خود نشان داده‌اند. نواحی کاندیدی که این دو روش ارائه داده‌اند، توسط شبکه RCNN بهترین نتیجه در برچسب‌گذاری اشیاء موجود از خود ارائه داده‌اند. در روش‌های پیشنهاد کاندیدا، تعداد ناحیه‌های پیشنهادی در این آزمایش، ۲۰۰۰ ناحیه لحاظ شده‌اند. باتوجه به نتایج به دست آمده روش R-CNN سریع با نواحی کاندیدا در اشکالی مانند بطری و گیاه نتایج خوبی از خود نشان نداده است، ولی در مواردی مانند قطار، اسب، گربه و ماشین این روش نتایج قابل قبولی را از خود نشان داده است.

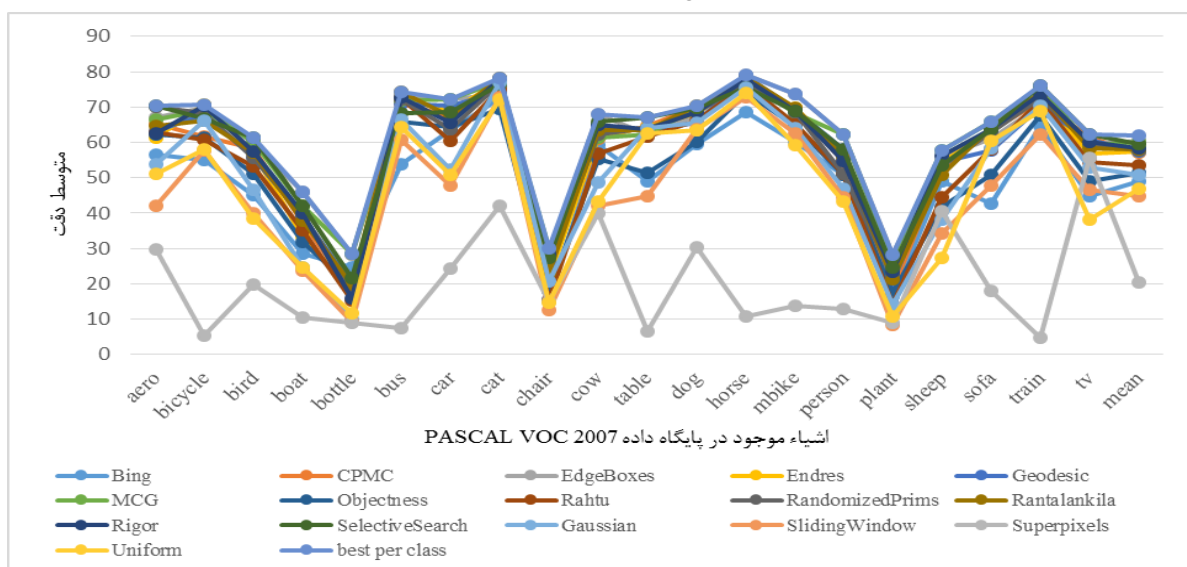
متفاوت بوده، از نظر آماری از لحاظ تعداد اشیای موجود در هر تصویر و اندازه اشیاء شبیه هم هستند.

مجموعه‌داده MS COCO شامل ۸۰ کلاس متفاوت از اشیاء است، ولی دارای اشیاء بیشتر و اندازه متفاوت‌تر (کوچک‌تر) در هر تصویر نسبت به دو مجموعه‌داده دیگر است.

## ۵. مقایسه بین روش‌های موجود

همان‌طور که در قسمت مقدمه بیان شد، یکی از چالش‌های اساسی و تأثیرگذار در بحث شناسایی اشیاء، پیشنهاد ناحیه‌های کاندیدا بوده است. از طرفی، یکی از روش‌های موفق و مشهور مطرح در زمینه شناسایی، روش R-CNN سریع بوده که توسط دانشگاه MIT پیاده‌سازی شده است. در همین راستا به بررسی تأثیر روش پیشنهاد کاندیدا بر روی این سیستم شناسایی پرداخته شده است. در آزمایش اول از مجموعه داده PASCAL VOC 2007 برای ارزیابی کلاس‌های موجود در این مجموعه داده به صورت جداگانه استفاده شده است. معیار ارزیابی استفاده شده در این مقاله، معیار دقت برچسب‌گذاری اشیاء بر پایه نواحی کاندیدی پیشنهادی است. در معیار دقت، نسبت برچسب‌های صحیح تشخیص داده‌شده اشیاء به تمامی اشیاء در نظر گرفته شده است.

برخی از روش‌های مطرح شده، نیاز به مرحله آموزش داشته‌اند. در همین راستا، در مجموعه‌داده‌های معرفی‌شده،



شکل (۱۱): مقایسه بین روش‌های پیشنهاد کاندیدا بر روی مجموعه‌داده PASCAL VOC 2007 با روش R-CNN سریع

(۲)، بار محاسباتی مناسب‌تری دارند و روش‌های Bing و EdgeBoxes بهترین عملکرد زمانی را دارند.

جدول (۲): مقایسه زمان متوسط اجرای روش‌های پیشنهاد کاندیدا بر

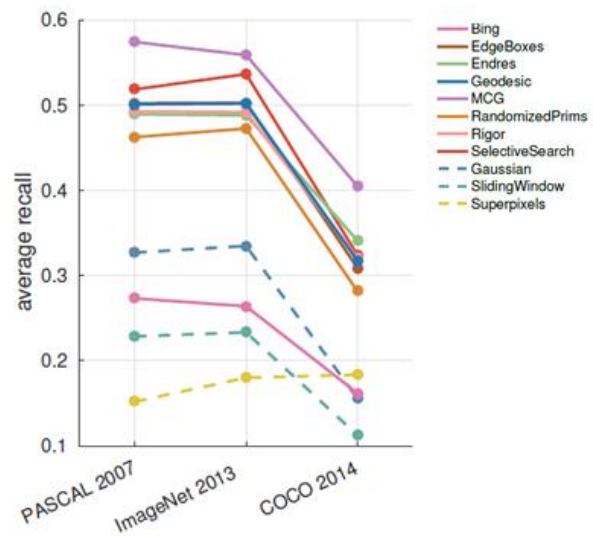
روی هر تصویر PASCAL VOC 2007

زمان بر حسب ثانیه	روش
۳	Objectness
۲۵۰	CPMC
۱۰۰	Endres
۱۰	Selective Search
۳	Rahtu
۱	Randomized Prims
۰٫۲	Bing
۳۰	MCG
۱۰	Ranta
۰٫۳	EdgeBoxes
۱	Superpixels

## ۶. نتیجه‌گیری

یکی از تنگناهای محاسباتی در سیستم‌های تشخیص و شناسایی اشیاء در تصویر، به‌دست‌آوردن ناحیه‌های کاندیدا هستند. در واقع عملکرد این نوع سیستم‌ها به روش انتخاب نواحی کاندیدا وابسته است. در همین راستا در این مقاله به معرفی و مقایسه روش‌های ناحیه‌های کاندیدا پرداخته شده و بررسی شد که این روش‌ها در یک مجموعه داده یکسان، چه عملکردی را از خود نشان می‌دهند. به‌منظور مقایسه این روش‌ها از یک سیستم شناسایی مشهور و محبوب R-CNN سریع که مورد استقبال بسیاری از محققان و پژوهشگران در این زمینه بوده، استفاده شده است. باتوجه به نتایج به‌دست‌آمده، استفاده از روش‌های پایه و مرجع به‌تنهایی عملکرد خوبی نداشته است و روش‌های EdgeBoxes و MCG بهترین نتایج را بر روی مجموعه‌داده‌های استاندارد از خود نشان داده‌اند.

در آزمایش دوم، از سه مجموعه‌داده استفاده شده است. شکل (۱۲) نتایج به‌دست‌آمده بر روی مجموعه PASCAL 2007، ImageNet 2013 و MS COCO را با هزار ناحیه کاندیدا را نشان می‌دهد. باتوجه به نتایج به‌دست‌آمده در هر سه مجموعه داده روش MCG بهترین نتایج را از خود نشان داده و روش‌های مرجع و پایه مانند ابرپیکسل، پنجره لغزان و گوسی به‌تنهایی نتایج قابل‌قبولی را از خود نشان نمی‌دهند.



شکل (۱۲): نتایج به‌دست‌آمده روی سه مجموعه‌داده متفاوت براساس ۱۰۰۰ ناحیه کاندیدا و روش‌های متفاوت کاندیدا

در ادامه برای ارزیابی بهتر روش‌های پیشنهاد کاندیدا، زمان اجرای این الگوریتم برای متوسط هر تصویر محاسبه شده است. این زمان‌ها در جدول (۲) آورده شده است. باتوجه به نتایج جدول (۲)، روش‌هایی مانند CPMC، Endres، جستجوی انتخابی، Ranta، MCG، بار محاسبات نسبتاً بالایی دارند و معمولاً برای شناسایی برخی اشیاء مناسب نیستند. ولی بقیه روش‌های ارائه‌شده در جدول

## مراجع

- [1] He, K., et al., *Spatial pyramid pooling in deep convolutional networks for visual recognition*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2015. 37(9): p. 1904-1916.
- [2] Ren, S., et al. *Faster R-CNN: Towards real-time object detection with region proposal networks*. in *Advances in Neural Information Processing Systems*. 2015.
- [3] Girshick, R., et al. *Rich feature hierarchies for accurate object detection and semantic segmentation*. in *Proceedings of the IEEE conference on computer vision and pattern*

- recognition. 2014.
- [4] Girshick, R. *Fast r-cnn*. in *Proceedings of the IEEE International Conference on Computer Vision*. 2015.
- [5] Simonyan, K. and A. Zisserman, *Very deep convolutional networks for large-scale image recognition*. arXiv preprint arXiv:1409.1556, 2014.
- [6] Uijlings, J.R., et al., *Selective search for object recognition*. International journal of computer vision, 2013. 104(2): p. 154-171.
- [7] Zitnick, C.L. and P. Dollár, *Edge boxes: Locating object proposals from edges*, in *Computer Vision–ECCV 2014*. 2014, Springer. p. 391-405.
- [8] Van de Sande, K.E., et al. *Segmentation as selective search for object recognition*. in *Computer Vision (ICCV), 2011 IEEE International Conference on*. 2011. IEEE.
- [9] Felzenszwalb, P.F. and D.P. Huttenlocher, *Efficient graph-based image segmentation*. International Journal of Computer Vision, 2004. 59(2): p. 167-181.
- [10] Van den Bergh, M., et al., *Seeds: Superpixels extracted via energy-driven sampling*. International Journal of Computer Vision, 2015. 111(3): p. 298-314.
- [11] Arbelaez, P., et al., *Contour detection and hierarchical image segmentation*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2011. 33(5): p. 898-916.
- [12] Dollár, P. and C.L. Zitnick, *Fast edge detection using structured forests*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2015. 37(8): p. 1558-1570.
- [13] Geusebroek, J.-M., et al., *Color invariance*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2001. 23(12): p. 1338-1350.
- [14] Manen, S., M. Guillaumin, and L. Gool. *Prime object proposals with randomized prim's algorithm*. in *Proceedings of the IEEE International Conference on Computer Vision*. 2013.
- [15] Rantalankila, P., J. Kannala, and E. Rahtu. *Generating object segmentation proposals using global and local search*. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014.
- [16] Chang, K.-Y., et al. *Fusing generic objectness and visual saliency for salient object detection*. in *Computer Vision (ICCV), 2011 IEEE International Conference on*. 2011. IEEE.
- [17] Carreira, J. and C. Sminchisescu. *Constrained parametric min-cuts for automatic object segmentation*. in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. 2010. IEEE.
- [18] Carreira, J. and C. Sminchisescu, *Cpmc: Automatic object segmentation using constrained parametric min-cuts*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2012. 34(7): p. 1312-1328.
- [19] Endres, I. and D. Hoiem, *Category-independent object proposals with diverse ranking*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2014. 36(2): p. 222-234.
- [20] Endres, I. and D. Hoiem, *Category independent object proposals*, in *Computer Vision–ECCV 2010*. 2010, Springer. p. 575-588.
- [21] Hoiem, D., A.A. Efros, and M. Hebert, *Recovering occlusion boundaries from an image*. International Journal of Computer Vision, 2011. 91(3): p. 328-346.
- [22] Krähenbühl, P. and V. Koltun, *Geodesic object proposals*, in *Computer Vision–ECCV 2014*. 2014, Springer. p. 725-739.
- [23] Humayun, A., F. Li, and J. Rehg. *RIGOR: reusing inference in graph cuts for generating object regions*. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014.
- [24] Lim, J., C. Zitnick, and P. Dollár. *Sketch tokens: A learned mid-level representation for contour and object detection*. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2013.
- [25] Yatziv, L., A. Bartesaghi, and G. Sapiro, *O (N) implementation of the fast marching algorithm*. Journal of computational physics, 2006. 212(2): p. 393-399.
- [26] Toivanen, P.J., *New geodesic distance transforms for gray-scale images*. Pattern Recognition Letters, 1996. 17(5): p. 437-450.
- [27] Bai, X. and G. Sapiro, *Geodesic matting: A framework for fast interactive image and video segmentation and matting*. International journal of computer vision, 2009. 82(2): p. 113-132.
- [28] Arbeláez, P., et al. *Multiscale combinatorial grouping*. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014.
- [29] Alexe, B., T. Deselaers, and V. Ferrari. *What is an object?* in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. 2010. IEEE.
- [30] Alexe, B., T. Deselaers, and V. Ferrari, *Measuring the objectness of image windows*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2012. 34(11): p. 2189-2202.
- [31] Rahtu, E., J. Kannala, and M. Blaschko. *Learning a category independent object detection cascade*. in *Computer Vision (ICCV), 2011 IEEE International Conference on*. 2011. IEEE.
- [32] Cheng, M.-M., et al. *BING: Binarized normed gradients for objectness estimation at 300fps*. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014.
- [33] Zhao, Q., Z. Liu, and B. Yin. *Cracking BING and*

- Beyond*. in *BMVC*. 2014.
- [34] Dollár, P. and C. Zitnick. *Structured forests for fast edge detection*. in *Proceedings of the IEEE International Conference on Computer Vision*. 2013.
- [35] Kim, J. and K. Grauman, *Shape sharing for object segmentation*, in *Computer Vision–ECCV 2012*. 2012, Springer. p. 444-458.
- [36] Erhan, D., et al. *Scalable object detection using deep neural networks*. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014.
- [37] Szegedy, C., et al., *Scalable, high-quality object detection*. arXiv preprint arXiv:1412.1441, 2014.
- [38] Yan, J., et al. *Object detection by labeling superpixels*. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015.