

بهبود الگوریتم تشخیص نقشه برجستگی مبتنی بر CRF با استفاده از ویژگی‌های مبتنی بر تجزیه ماتریس*

محمد شوریابی^۱

محمد جواد فدائی اسلام^۲

چکیده

با توجه به این که نقشه برجستگی تصویر می‌تواند در الگوریتم‌های مانند بخش‌بندی، فشرده‌سازی و بازیابی تصویر کاربرد داشته باشد، ارائه یک مدل کارآمد برای تشخیص برجستگی مورد توجه محققین قرار گرفته است. علیرغم این که تاکنون کارهای زیادی در این زمینه انجام شده است، اما تاکنون یک مدل مؤثر و کارآمد که بتواند با محاسبات کم نقشه برجستگی تصویر را آشکار کند، ارائه نشده است. برای این منظور، ما یک الگوریتم ساده و تحت نظارت برای شناسایی نقشه برجستگی با استفاده از میدان تصادفی شرطی (CRF) و نشانه‌های برجستگی پیشنهاد می‌کنیم. در روش پیشنهادی برای آموزش CRF از ویژگی‌های کنتراست محلی، مرکز سوگیری و پس‌زمینه‌ای استفاده شده است، علاوه بر این سه ویژگی برای کارایی بهتر، ویژگی جدیدی مبتنی بر تجزیه ماتریس به کار گرفته شده است. در ادامه CRF با توجه به ویژگی‌های ۲۰ تصویر که به تصویر ورودی نزدیک هستند، آموزش می‌بیند. در نهایت برجستگی تصویر ورودی با توجه به وزن‌های محاسبه شده در مرحله آموزش، نشانه‌های برجستگی تصویر ورودی و مبنای درستی محاسبه می‌شود. روش پیشنهادی در دقت و سرعت اجرای الگوریتم نسبت به سایر روش‌ها برتری دارد.

کلمات کلیدی: ابر پیکسل، تشخیص برجستگی، میدان تصادفی شرطی، نشانه‌های برجستگی.

* تاریخ دریافت: ۹۷/۹/۱۴؛ تاریخ پذیرش: ۹۷/۱۰/۲۶.

^۱ دانشجوی کارشناسی ارشد مهندسی کامپیوتر، دانشکده برق و کامپیوتر، دانشگاه سمنان، سمنان، ایران (نویسنده مسئول)

m.shouryabi@semnan.ac.ir

fadaei@semnan.ac.ir

^۲ استادیار گروه مهندسی کامپیوتر، دانشکده برق و کامپیوتر، دانشگاه سمنان، سمنان، ایران

مقدمه

صحنه‌هایی که ما در زندگی روزمره با آن‌ها مواجه می‌شویم، معمولاً توسط عوامل مختلفی دستخوش تغییر می‌شوند. ظرفیت سیستم بینایی برای پردازش این تغییرات و اطلاعات صحنه در هر لحظه محدود است. توجه بینایی مکانیزی است که نواحی مهم و چشم‌گیر را در یک صحنه مشخص کرده، آن‌ها را برای پردازش‌های سطح بالاتر همچون شناسایی هدف آماده می‌کند. در حقیقت توجه بینایی باعث شده تا انسان در محیط اطرافش بی‌درنگ به جستجو و شناسایی اشیاء مختلف و مورد علاقه‌اش بپردازد. توجه بینایی را می‌توان یکی از ابزارهای ابتدایی برای طراحی سیستم مصنوعی چشم انسان در نظر گرفت به گونه‌ای که توسط این سازوکار می‌توان نواحی مورد توجه را نسبت به بقیه نواحی با وضوح بیشتری مشاهده نمود و نواحی اطراف آن را به عنوان ورودی‌های غیرضروری در درک تصویر بیان کرد (رولز و دکو، ۲۰۰۶). توجه بینایی به ویژه در ۲۵ سال اخیر بسیار مورد اهمیت قرار گرفته و مدل‌های متنوعی از آن نیز در حوزه‌های مختلف ارائه شده است. محققین زیادی جهت ارائه یک مدل محاسباتی برای طراحی این سازوکار تلاش کرده‌اند. این مدل‌ها بسته به ویژگی‌های به کار گرفته در آن‌ها به دو نوع بالا به پایین^۱ و پایین به بالا^۲ معروفند. مدل‌های بالا به پایین هدف محور می‌باشند که دارای پردازش اطلاعات سطح بالای تصاویر که برگرفته از دانش قبلی در رابطه با صحنه‌اند مانند بازشناسی شی، طبقه‌بندی صحنه و شناسایی هدف می‌باشند. در سوی دیگر، توجه پایین به بالا برگرفته از ویژگی‌های اولیه تصویر مانند رنگ، روشنایی و جهت می‌باشد که به آن پردازش محرک محور می‌گویند.

نگاه رایج این است که توجهات بالا به پایین و محرک محور، با یکدیگر ترکیب شوند تا رفتار توجهی ما را هدایت کنند. یک مدل و روش ترکیبی بایستی قادر به توضیح این امر باشد که چه موقع و چگونه به یک مورد بینایی محیط اطراف توجه شده، یا به علت وجود علائم برجسته یک عامل محرک دیگر، از آن صرف‌نظر می‌شود. بیشتر مدل‌های

¹ Top-down² Bottom-up

توجه، وابسته به عامل پایین به بالا است و واژه نواحی برجسته^۱ به این مدل‌ها اطلاق می‌شود و وجود اکثریت، در این حوزه به دلیل وجود روش‌های استاندارد و سادگی پیاده‌سازی آن‌ها می‌باشد (برجی و ایتی، ۲۰۱۳).

طی دهه‌های اخیر، محققان پژوهش‌های زیادی پیرامون تعامل مکانیزم توجه بینایی و مدل‌های محاسباتی آن انجام داده‌اند. تحقیقات تشخیص برجستگی در پیش از سال‌های ۲۰۱۰، در مرحله پیشرفت‌های بنیادی قرار دارند و توجه رشته‌های مختلف از جمله روانشناسی شناختی، علوم اعصاب و بینایی کامپیوتری را به خود جلب کرده است. شاید اولین تلاش توسط Treisman در سال ۱۹۸۰ بود که منجر به پیدایش نظریه یکپارچگی ویژگی شد. بر اساس این نظریه، اکثر نواحی برجسته صحنه‌های طبیعی در زمانی که چشم مشاهده‌گر به صحنه دوخته شده است در ویژگی‌های سطح پایین نظیر رنگ، روشنایی و جهت نسبت به بقیه نواحی دارای تمایز می‌باشند (اکاتی و ابراهیمی مقدم و هراتی، ۱۳۹۴). به دنبال این نظریه، Itti و همکاران (ایتی و کوچ و نیبور، ۱۹۹۸) یک مدل مرکز-پیرامون چند مقیاسی را ارائه می‌کنند که در آن به منظور تشخیص برجستگی، رنگ، شدت و جهت‌گیری با مقیاس‌های مختلف با هم ادغام می‌شوند. Bruce و همکاران (براک و تسوتوس، ۲۰۰۶) برای محاسبه برجستگی از اندازه‌گیری مقدار اطلاعاتی شانون در زمینه محلی استفاده می‌کنند. (آچانتا و حمامی و استرادا و استرانک، ۲۰۰۹)، روش فرکانس تنظیم شده مبتنی بر رنگ و درخشندگی را ارائه می‌نمایند، در این روش مقدار برجستگی توسط تفاوت رنگ و با توجه به میانگین ارزش پیکسلی محاسبه می‌شود. (هو و ژانگ، ۲۰۰۷) و (گو و ما و ژانگ، ۲۰۰۸) برای تشخیص برجستگی بصری از تجزیه و تحلیل طیف فوریه استفاده می‌کنند.

از سال ۲۰۱۰، به منظور تشخیص برجستگی، مدل‌های پیشرفته‌تر و به ویژه مدل‌های مبتنی بر گراف مطرح شده‌اند که به طور کلی دقت کلی تشخیص را بهبود می‌بخشند. (ژیانگ و ژانگ و لو و یانگ و یانگ، ۲۰۱۳) یک مدل گراف دو حلقه‌ای ایجاد نموده‌اند

¹ Salient Region

که در آن مقادیر برجستگی در مناطق مختلف تصویر با احتمالات جذب مارکوف محاسبه می‌شود. (لی، لی، وانگ، ایرل و فنگ، ۲۰۱۳)، برای غلبه بر تأثیر منفی الگوهای تصویر با مقیاس کوچک و کنتراست بالا، یک رویکرد چند لایه را پیشنهاد کردند که به وسیله مدل درختی سلسله مراتبی موجب بهبود تشخیص برجستگی می‌شود. (پرازی و کراهنبل و پریچ و هورنونگ، ۲۰۱۲)، محاسبات کنتراست و برجستگی را در چارچوب فیلترینگ گاوسی با ابعاد بالا هماهنگ می‌کنند. (وی و ون و ژو و سان، ۲۰۱۲) از دانش اولیه پس زمینه‌ها و فاصله جغرافیایی برای محاسبه برجستگی بصری استفاده می‌نمایند. (یانگ و ژانگ و لو و روان و یانگ، ۲۰۱۳) برای استخراج نمایش‌های پیش زمینه در نقشه برجستگی نهایی، از رتبه‌بندی چند منظوره مبتنی بر گراف استفاده می‌کنند و در این رتبه‌بندی از چهار مرز تصویر به عنوان دانش اولیه پس زمینه استفاده می‌شود.

اگر شروع بررسی‌ها در مورد برجستگی تصویر را سال ۱۹۸۰ در نظر بگیریم، هر چه به سمت جلو پیش می‌رویم، معمولاً برای تشخیص برجستگی تصویر از رویکردهای یادگیری تحت نظارت استفاده شده است. در مطالعه (یانگ و یانگ، ۲۰۱۷) در تمایز برجستگی بصری از یادگیری مشترک میدان تصادفی شرطی^۱ (CRF) استفاده شده است. (لو و ماهادوان و وسکونکوس، ۲۰۱۴)، فرآیند انتشار مبتنی بر گراف را برای یادگیری بذره‌های بهینه^۲ تصویر و تمایز میان شی و پس زمینه اعمال کردند. (لو و ماهادوان و وسکونکوس، ۲۰۱۴)، برای تجمیع نقشه‌های برجستگی مدل‌های مختلف، مدل CRF را آموزش می‌دهند که نه تنها از نقشه‌های برجستگی منحصر به فرد بلکه از تعاملات بین پیکسل‌های مختلف نیز بهره می‌برد. در مطالعه (تونگ و لو و روان و یانگ، ۲۰۱۵)، از چند نمونه از نقشه برجستگی ضعیف به عنوان مجموعه آموزشی ماشین‌های بردار منبع^۳ (SVM) و متعاقباً از آن‌ها برای ایجاد یک نقشه برجسته قوی استفاده می‌کنند.

¹ Conditional Random Fields

² optimal seeds

³ supply vector machines

پیشینه پژوهش

در بخش حاضر به معرفی دو مدل که در روش پیشنهادی مورد استفاده قرار گرفته‌اند، می‌پردازیم.

رویکرد CRF مبتنی بر سوپر پیکسل برای تشخیص برجستگی

(کویو و گاو و هان، ۲۰۱۷) یک الگوریتم ساده و جدید تحت نظارت برای شناسایی نقشه برجستگی با استفاده از میدان‌های تصادفی شرطی و نشانه‌های برجستگی (کنتراست محلی^۱، مرکز سوگیری^۲ و پس‌زمینه‌ای^۳) پیشنهاد می‌کنند. آن‌ها معتقداند بیشتر رویکردهای CRF موجود، مدل‌های گرافیکی احتمالی را با گراف پیکسل همسایگی ۸-گانه شبکه‌ای شکل تنظیم می‌کنند، در حالی که بررسی گراف سطح ابرپیکسل آن‌ها نه تنها می‌تواند مدل را ساده کند، بلکه باعث می‌شود هر پیچ ابرپیکسل، همسایگی و اطلاعات مرزی بیشتری داشته باشد که برای یادگیری و نتیجه‌گیری در مدل گراف مفید است. علاوه بر این، به ظاهر عنصر ابرپیکسل برای سیستم بصری انسان مناسب‌تر است. اساس کار این مقاله به شرح زیر است:

- برای تشخیص منطقه برجسته، یک گراف مبتنی بر ابرپیکسل برای میدان تصادفی شرطی ایجاد می‌شود.
- در رابطه با هر تصویر ورودی، تعدادی تصویر که نزدیک به تصویر ورودی می‌باشد، توسط الگوریتم gist (اولیوا و تورالبا، ۲۰۰۱) انتخاب می‌شود، سپس یک مدل میدان تصادفی شرطی خاص توسط این تصاویر که به تصویر ورودی نزدیک هستند، آموزش می‌بیند.
- (اولیوا و تورالبا، ۲۰۰۱) در بخش اول از الگوریتم SLIC برای تجزیه به ابرپیکسل استفاده می‌کنند^۴. اما در بخش دوم با توجه به این که به طور کلی شبکه چشم انسان با دو

نوع سلول بصری شناخته شده و پر می شود: یعنی مخروطی و میله‌ای و سلول‌های مخروطی می‌توانند به انسان کمک کنند که یک دید با جزئیات در مورد نواحی مورد علاقه به دست آورد و سلول‌های میله‌ای یک دید کلی را در چندین میلی ثانیه اول ارائه می‌دهند و پس از آن انسان می‌تواند توسط تجربه بدون هیچ درنگی برترین ناحیه جذاب را برآورد کند. بنابراین با توجه به مواردی که بیان شد (اولیوا و تورالبا، ۲۰۰۱) معتقد بودند که علاوه بر دید اولیه، مسئله تجربه را هم باید به گونه‌ای در روش پیشنهادی خود وارد کنند. برای این منظور یک سری کاندیدا را که صحنه مشابهی با تصویر پرس جو دارند را با استفاده از ویژگی gist پیدا می‌کنند. کاندیداهای انتخاب شده را برای هر تصویر برای آموزش یک مدل مجزا وارد CRF می‌کنند.

تشخیص شیء برجسته از طریق تجزیه ماتریس ساختار یافته

(پنگ و لی و لیانگ و هو و ژیونگ، ۲۰۱۷) یک روش جدید برای تجزیه ماتریس‌ها به منظور یافتن نواحی برجسته ابداع نموده‌اند. (پنگ و لی و لیانگ و هو و ژیونگ، ۲۰۱۷) در ابتدا با توجه به روش SLIC تصویر را به ابرپیکسل تبدیل می‌کنند. سپس از هر ابرپیکسل ۵۳ ویژگی استخراج می‌کنند و ماتریس F را با ابعاد $N \times 53$ (تعداد ابرپیکسل‌هاست) تشکیل می‌دهند. سپس ساختار درختی را به همراه اطلاعات اولیه‌ای از میزان برجستگی نواحی به مسئله اضافه می‌کنند. (پنگ و لی و لیانگ و هو و ژیونگ، ۲۰۱۷) در این مقاله سعی دارند موضوع رتبه پایین، اطلاعات مربوط به همسایگی و اطلاعات اولیه مربوط به نقشه برجستگی را به نوعی با هم ترکیب نمایند. آن‌ها از رابطه (۱) برای انجام موارد بیان شده (موضوع رتبه پایین، اطلاعات همسایگی و اطلاعات اولیه نقشه برجستگی) بهره می‌گیرند. در این مقاله، نقشه برجستگی تصویر (F) از تجزیه ماتریسی به L (ماتریس مرتبه پایین) که نشان‌دهنده پیش زمینه تصویر و S (ماتریس پراکنده) که نشان‌دهنده شیء برجسته می‌باشد، محاسبه می‌شود (شکل ۱).

$$\min_{L,S} \psi(L) + \alpha\Omega(S) + \beta\theta(L,S)$$

$$\text{s. t.} \quad \text{رابطه ۱)} \\ F = L + S$$

در این رابطه Ψ نشان‌دهنده رتبه ماتریس L ، تابع Ω در بردارنده یک ساختار درختی است که در آن اطلاعاتی از میزان اهمیت هر ناحیه وجود دارد و Θ یا تنظیم لاپلاسی باعث افزایش شکاف بین اشیا برجسته و پس‌زمینه در فضای ویژگی می‌شود.

روش پژوهش

روش میدان تصادفی شرطی در سال ۲۰۰۱ توسط (لافرتی و مک کالم و پیرا، ۲۰۰۱) ارائه شد، CRFها مدل‌های احتمالی برای برجسب‌زنی داده‌های متوالی می‌باشند (لافرتی و مک کالم و پیرا، ۲۰۰۱). میدان تصادفی شرطی که ما در این مقاله پیشنهاد می‌کنیم، برگرفته از مقاله (کویو و گاو و هان، ۲۰۱۷) می‌باشد، CRF پیشنهادی در واقع مدلی احتمالی برای محاسبه $P(A|F)$ می‌باشد که در آن $A = \{a_x\}$ مجموعه برجسب‌ها و $F = \{f_1, f_2, f_3\}$ نشانه‌های برجستگی می‌باشد. CRF پیشنهاد شده به صورت معادله (۲) تعریف می‌شود.

$$P(A|F) = \frac{1}{Z} \exp\{-E(A|F)\} \quad \text{رابطه ۲}$$

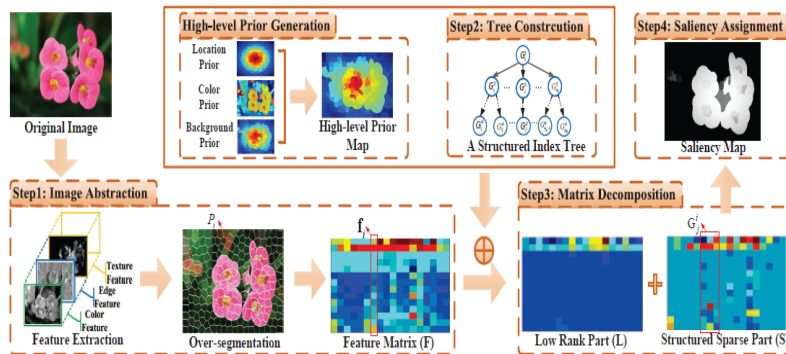
که در آن Z ثابت نرمال‌سازی به نام تابع افراز و $E(A|F)$ تابع انرژی می‌باشد. (کویو و گاو و هان، ۲۰۱۷) سعی کردند نشانه‌های برجستگی و سیستم همسایگی را در تابع انرژی مورد نظر ایجاد کنند. بنابراین تابع انرژی را به صورت صورت رابطه (۳) تعریف کردند.

$$E(E|A) = \sum_{x \in S} \sum_{i=1}^m \lambda_i V_i(a_x, f_i) + \sum_{x \in S, x' \in N(x)} N(a_x, a_{x'}, F) \quad \text{رابطه ۳}$$

که در آن V_i و N به ترتیب پتانسیل وابستگی و اثر متقابل (تعامل) نامیده می‌شوند. به‌طور ویژه، پتانسیل وابستگی در ناحیه x ، تابعی از تمام ویژگی مشاهده شده (نشانه برجستگی) F و همین‌طور برجسب A می‌باشد. در واقع این تابع تأثیر نشانه‌های برجستگی را روی برجسب برجستگی (برجسته بودن یا نبودن) نشان می‌دهد. (کویو و گاو و هان، ۲۰۱۷) پتانسیل وابستگی را به صورت $V_i(a_x, f_i) = f_{ix} a_x$ تشکیل می‌دهند که در آن مقدار نشانه برجستگی i در ابریکسل x و a_x برابر مبنای درستی برجستگی می‌باشد. پتانسیل وابستگی فقط به گره (ناحیه x) وابسته است در حالی که پتانسیل تعامل ارتباط بین دو ناحیه x و همسایه آن x' را در نظر می‌گیرد. پتانسیل تعامل توسط (کویو و گاو و هان، ۲۰۱۷) به صورت رابطه (۴) تعریف می‌شود.

$$N(a_x, a_{x'}, F) = 1(a_x \neq a_{x'}) \exp(-\beta \|F_x - F_{x'}\|) \quad \text{رابطه ۴}$$

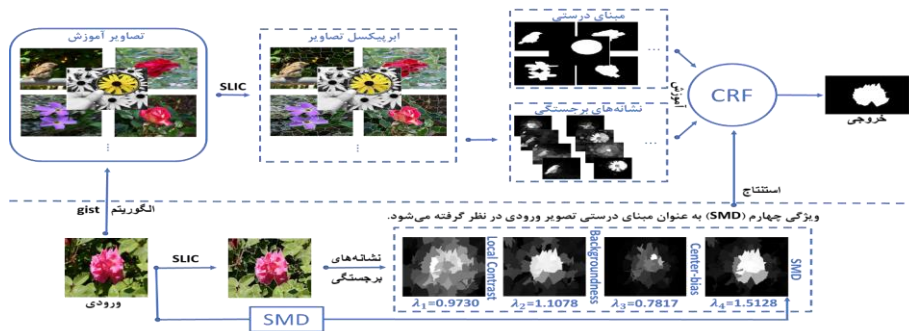
λ_i موجود در رابطه (۳) وزن (ضریب تأثیر) هر نشانه برجستگی را نشان می‌دهد. مقادیر λ_i با توجه به داده‌های آموزش محاسبه می‌شوند. (کویو و گاو و هان، ۲۰۱۷) برای قسمت آموزش CRF، تعداد ۲۰ تصویر را که نزدیک به تصویر ورودی می‌باشد را از کل دیتاست توسط الگوریتم gist به دست می‌آورند. سپس هر کدام از تصاویر توسط الگوریتم SLIC ناحیه‌بندی می‌شوند. تفاوت روش ما با روش پیشنهادی (کویو و گاو و هان، ۲۰۱۷) در این است که آن‌ها از سه نشانه‌ی برجستگی کنتراست محلی، مرکز سوگیری و پس‌زمینه‌ای برای هر ناحیه از تصاویر آموزش استفاده می‌کنند، در حالی که ما علاوه بر نشانه‌های برجستگی بیان شده خروجی روش پیشنهادی Peng و همکاران را به عنوان ویژگی چهارم به سیستم مورد نظر می‌دهیم. با توجه به ۲۰ تصویری که برای آموزش در نظر گرفته شده و موجود بودن مبنای درستی و نشانه‌های برجستگی هر تصویر، مقادیر λ_i برای هر نشانه برجستگی مشخص می‌شود.



شکل ۱. نمایشی از روش پیشنهادی در (بنگ و لی و لیانگ و هو و ژیونگ، ۲۰۱۷)

هر نشانه برجستگی که بهتر باشد، ضریب λ مربوط به آن بیشتر خواهد بود. پس از محاسبه λ ها، مشخص می‌شود هر ویژگی تا چه اندازه برای تشخیص ناحیه برجسته مفید بوده است. باید به این نکته توجه داشت که در مرحله آموزش مقدار مبنای درستی هر تصویر مشخص است اما در مرحله آزمایش مقدار مبنای درستی تصویر ورودی مشخص نیست، Qiu و همکاران مقدار a_x را به صورت تصادفی تولید می‌کنند، اما پیشنهادی که ما

برای این قسمت از الگوریتم داریم، این است که با توجه به این که CRF به هر یک از نشانه‌های برجستگی موجود وزنی را اختصاص می‌دهد، بنابراین ما ویژگی که وزن بیشتری دارد را به عنوان مبنای درستی تصویر ورودی در نظر بگیریم. در قسمت استنتاج با توجه به وزن‌های محاسبه شده در قسمت آموزش، نشانه‌های برجستگی موجود برای تصویر ورودی و همچنین مقدار مبنای درستی که نحوه محاسبه آن ذکر شد، برجستگی تصویر ورودی محاسبه می‌شود (شکل ۲).



شکل ۲. چارچوب مدل برجستگی پیشنهادی

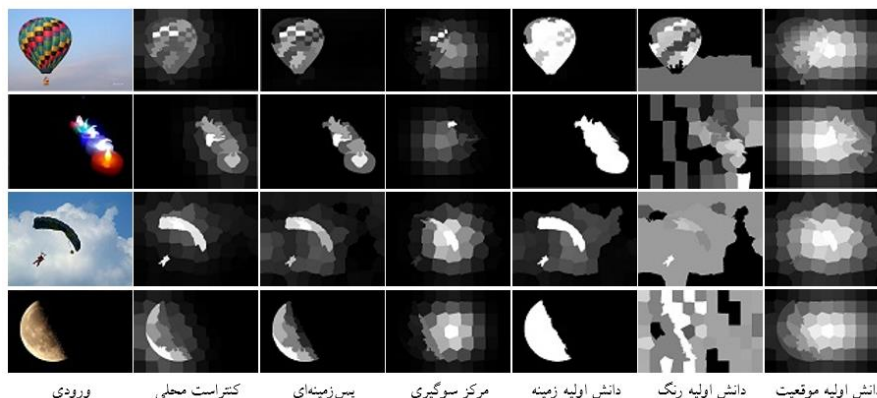
در ادامه این کار ما سعی کردیم، روش پیشنهادی را با ترکیب ویژگی‌های دو مقاله Qiu و همکاران (کنتراست محلی، مرکز سوگیری و پس‌زمینه‌ای) و Peng و همکاران (دانش اولیه موقعیت^۱، دانش اولیه رنگ^۲، دانش اولیه زمینه^۳) بهبود ببخشیم (شکل ۳). به طور کلی، ترکیب ویژگی‌های مختلف برای استفاده در یک الگوریتم که همه از آن بهره‌مند شوند، مشکل است. اگرچه چند آزمایش یکپارچه‌سازی انجام شده است (ژیانگ و وانگ و یوان و لیو و ژنگ، ۲۰۱۳)، (فو و گونگ و یانگ و ژو و گو، ۲۰۱۳)، اما این الگوریتم‌ها به طور خاص، با توجه به دارا بودن تعداد زیادی از ویژگی‌ها، بسیار بزرگ و ناکارآمد طراحی شده‌اند. ما ترکیب‌های مختلف ویژگی‌ها را در روش پیشنهادی بررسی کردیم و بهترین عملکرد در استفاده از چهار ویژگی کنتراست محلی، پس‌زمینه‌ای، دانش

¹ Location Prior

² Color Prior

³ Background Prior

اولیه موقعیت و زمینه به دست آمد (روش پیشنهادی ۲). اما نتایج حاصل از این ترکیب به خوبی نتایج روش پیشنهادی اول نیست.



شکل ۳. ویژگی‌های مورد استفاده در (اویو و گاو و هان، ۲۰۱۷) و (پنگ و لی و لیانگ و هو و ژیونگ، ۲۰۱۷)

نتایج پیاده‌سازی

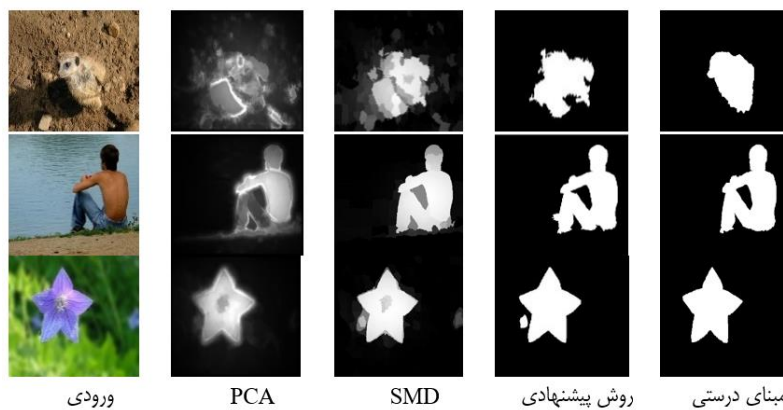
برای نشان دادن اثربخشی روش تشخیص برجستگی پیشنهادی، روش خود را با روش‌های تشخیص برجستگی مختلف مقایسه می‌کنیم. مدل پیشنهادی با مدل‌های تشخیص برجستگی PCA (مارگولین و تال و ژنلیک مانور، ۲۰۱۳)، CB (ژیانگ و وانگ و یوان و لیو و ژنگ لی، ۲۰۱۱)، wCtr (ژو و لیانگ و وی و سان، ۲۰۱۴)، SMD (پنگ و لی و لیانگ و هو و ژیونگ، ۲۰۱۷) و همچنین روش پیشنهادی Qiu و همکاران که اساس کار ما هم همین روش می‌باشد، بر روی پایگاه داده MSRA-10K که شامل ۱۰۰۰۰ تصویر است و اخیراً در پژوهش‌های فراوانی استفاده شده است، مقایسه می‌گردد. مقایسه بصری نقشه‌های برجستگی مدل پیشنهادی نسبت به مدل‌های دیگر در شکل ۴ نشان داده شده است. همان‌طور که در شکل مشاهده می‌کنید، نقشه برجستگی حاصل از مدل پیشنهادی ما، به مبنای درستی تصویر مورد نظر نزدیک‌تر می‌باشد اما به منظور ارزیابی دقیق‌تر روش پیشنهادی به سراغ معیارهای ارزیابی می‌رویم.

روش‌های مختلفی برای ارزیابی کمی نقشه‌های برجستگی پیشنهاد شده‌اند. اساس ارزیابی باید نزدیکی بین نقشه برجستگی مشتق شده و نقشه مرجع را تعیین کند. در این مقاله برای ارزیابی روش پیشنهادی از معیارهای میانگین خطای مطلق (MAE) و F-measure استفاده می‌کنیم. معیار F-measure از مقادیر دقت و حساسیت محاسبه می‌شود که در پژوهش‌های اخیر، زیاد مورد استفاده قرار گرفته است. در اغلب موارد معیارهای دقت و حساسیت ارزیابی جامعی از کیفیت نقشه برجستگی ارائه نمی‌کنند. به همین دلیل، معیار F-measure میانگین وزنی دو معیار دقت و حساسیت با وزن نامنفی β^2 پیشنهاد می‌شود.

$$F_{\beta} = \frac{(1 + \beta^2) \times \text{Precision} \times \text{Recall}}{\beta^2 \times \text{Precision} + \text{Recall}} \quad \text{رابطه (۵)}$$

که در آن β^2 پیرو [۲۷-۳۰]، مقدار ۰/۳ تنظیم می‌شود. آستانه طوری تعیین می‌شود که منجر به بهترین F_{β} می‌شود. معیار F-measure تأثیر تعیین پیکسل‌های غیربرجسته و همچنین اهمیت تشخیص کامل را در نظر نمی‌گیرد. به همین دلیل برای پرداختن به این مسائل از معیار MAE استفاده می‌کنیم (ژاو و اوایانگ و لی و وانگ، ۲۰۱۵). با توجه به نقشه برجستگی S و مبنای درستی باینری G، MAE به عنوان میانگین خطای قدر مطلق بین S و G تعریف می‌شود (پنگ و لی و لیانگ و هو و ژیونگ، ۲۰۱۷):

$$\text{MAE} = \text{mean}(|S - G|) \quad \text{رابطه (۶)}$$



شکل ۴. مقایسه بصری مدل پیشنهادی با سایر مدل‌ها

با توجه به معیارهای که در این بخش بیان شد، همانطور که در جدول ۱ مشاهده می‌کنید، روش پیشنهادی اول نسبت به سایر روش‌ها بهتر عمل کرده است و این حاکی از کارایی مناسب آن نسبت به سایر روش‌ها می‌باشد. برای مقایسه روش پیشنهادی خود از نظر زمان اجرا با سایر روش‌ها به خصوص روش پیشنهادی (کویو و گاو و هان، ۲۰۱۷) که اساس کار ما نیز این روش می‌باشد، ما با توجه به این که کد مربوط به روش پیشنهادی (کویو و گاو و هان، ۲۰۱۷) را در اختیار نداشتیم، تصمیم گرفتیم میانگین زمانی روش پیشنهادی خود را با روش‌های که (کویو و گاو و هان، ۲۰۱۷) با روش پیشنهادیشان مقایسه کرده بودند، مقایسه کنیم. جدول ۲ میانگین زمانی اجرای الگوریتم برای هر تصویر را در سیستمی با مشخصات Intel(R) Xeon(R) CPU E5640 @ 2.67GHz و 20.00 GB (RAM) نشان می‌دهد و جدول ۳ میانگین زمانی می‌باشد که Qiu و همکاران برای هر تصویر در مقاله خود بیان کرده‌اند. با مقایسه این دو جدول به راحتی می‌توان متوجه شد که حتی اگر سیستم مورد استفاده توسط ما و (کویو و گاو و هان، ۲۰۱۷) را یکسان در نظر بگیریم، الگوریتم پیشنهادی ما با وجود کارایی مناسب، زمان اجرای پایین‌تری را نیز دارا می‌باشد. اما همان‌طور که واضح است به راحتی با مقایسه جدول ۲ و ۳ می‌توان متوجه شد که سیستم مورد استفاده ما ضعیف‌تر می‌باشد. بنابراین، با وجود این که ما علاوه بر ویژگی‌های (کویو و گاو و هان، ۲۰۱۷) ویژگی چهارمی را هم به مدل پیشنهادی اضافه کردیم، اما باز هم به زمان اجرا بسیار مناسبی نسبت به مقاله پایه خود دست پیدا کرده‌ایم.

جدول ۱. معیار MAE و WfB برای هر مدل

روش‌ها	MAE	WfB
روش پیشنهادی ۱	۰/۰۸۱۴	۰/۷۹۸۳
روش پیشنهادی ۲	۰/۰۹۸۲	۰/۷۰۰۲
SMD	۰/۱۰۴۶	۰/۷۰۵۰
Qiu و همکاران	۰/۰۸۶۰	۰/۷۰۹۵
CB	۰/۲۰۸۱	۰/۴۳۲۶
PCA	۰/۱۸۵۳	۰/۴۸۳۰
wCtr	۰/۱۲۱۱	۰/۶۶۵۴

جدول ۲. میانگین زمان اجرای روش‌های مختلف

روش‌ها	زمان (ثانیه)
روش پیشنهادی ۱	۰/۹۳
روش پیشنهادی ۲	۵/۴۶
SMD	۳/۲۸
PCA	۶/۰۰
CB	۲/۶۰
wCtr	۰/۵۱۲

جدول ۳. میانگین زمان اجرای روش‌های مختلف توسط Qiu و همکاران

روش‌ها	زمان (ثانیه)
Qiu و همکاران	۱۷/۱۳
PCA	۲/۳۴
CB	۱/۷۹
wCtr	۰/۳۱

نتیجه‌گیری و پیشنهادها

در این مقاله با توجه به روش پیشنهادی (کویو و گاو و هان، ۲۰۱۷) و همچنین Peng و همکاران، روش جدیدی برای شناسایی نواحی برجسته ارائه شد که در این روش، CRF با استفاده از ویژگی‌های کنتراست محلی، مرکز سوگیری، پس‌زمینه‌ای و خروجی SMD آموزش می‌بیند، در قسمت آموزش از ۲۰ تصویر که نزدیک به تصویر ورودی می‌باشد، استفاده شده است. برای آموزش CRF اطلاعات ۲۰ تصویر که این اطلاعات شامل ویژگی‌ها (نشانه‌های برجستگی) و مبنای درستی می‌باشد، استفاده شد. سپس با توجه به وزن‌های محاسبه شده در قسمت آموزش، نشانه‌های برجستگی موجود برای تصویر ورودی و همچنین مقدار مبنای درستی که نحوه محاسبه آن ذکر شد، برجستگی تصویر ورودی محاسبه شد. در ادامه سعی شد ترکیبی از ویژگی‌های مختلف به عنوان نشانه‌های برجستگی در روش پیشنهادی استفاده شود. الگوریتم پیشنهادی به وسیله‌ی معیارهای ارزیابی MAE و Wfb و در پایگاه MSRA-10K با روش‌های دیگر مقایسه شد. نتایج حاصله بیانگر دقت و

۱۶۴ // دوفصلنامه مدیریت مهندسی و رایانش نرم

سرعت مناسب روش پیشنهادی نسبت به سایر روش‌ها می‌باشد.

منابع

- Rolls, E. T., & Deco, G. (2006). Attention in natural scenes: neurophysiological and computational bases. *Neural networks*, 19(9), 1383-1394.
- Borji, A., & Itti, L. (2013). State-of-the-art in visual attention modeling. *IEEE transactions on pattern analysis and machine intelligence*, 35(1), 185-207.
- Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (11), 1254-1259.
- Bruce, N., & Tsotsos, J. (2006). Saliency based on information maximization. In *Advances in neural information processing systems* (pp. 155-162).
- Achanta, R., Hemami, S., Estrada, F., & Süsstrunk, S. (2009). Frequency-tuned salient region detection. In *IEEE international conference on computer vision and pattern recognition (CVPR 2009)* (No. CONF, pp. 1597-1604).
- Hou, X., & Zhang, L. (2007, June). Saliency detection: A spectral residual approach. In *2007 IEEE Conference on computer vision and pattern recognition* (pp. 1-8). IEEE.
- Guo, C., Ma, Q., & Zhang, L. (2008, June). Spatio-temporal saliency detection using phase spectrum of quaternion fourier transform. In *2008 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1-8). IEEE.
- Jiang, B., Zhang, L., Lu, H., Yang, C., & Yang, M. H. (2013). Saliency detection via absorbing markov chain. In *Proceedings of the IEEE international conference on computer vision* (pp. 1665-1672).
- Li, A., Li, C., Wang, X., Eberl, S., Feng, D. D., & Fulham, M. (2013, November). Automated segmentation of prostate MR images using prior knowledge enhanced random walker. In *2013 International Conference on Digital Image Computing: Techniques and Applications (DICTA)* (pp. 1-7). IEEE.
- Perazzi, F., Krähenbühl, P., Pritch, Y., & Hornung, A. (2012, June). Saliency filters: Contrast based filtering for salient region detection. In *2012 IEEE conference on computer vision and pattern recognition* (pp. 733-740). IEEE.
- Wei, Y., Wen, F., Zhu, W., & Sun, J. (2012, October). Geodesic saliency using background priors. In *European conference on computer vision* (pp. 29-42). Springer, Berlin, Heidelberg.
- Yang, C., Zhang, L., Lu, H., Ruan, X., & Yang, M. H. (2013). Saliency detection via graph-based manifold ranking. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3166-3173).
- Yang, J., & Yang, M. H. (2017). Top-down visual saliency via joint CRF and dictionary learning. *IEEE transactions on pattern analysis and machine intelligence*, 39(3), 576-588.
- Lu, S., Mahadevan, V., & Vasconcelos, N. (2014). Learning optimal seeds for diffusion-based salient object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2790-2797).
- Lu, S., Mahadevan, V., & Vasconcelos, N. (2014). Learning optimal seeds for diffusion-based salient object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2790-2797).
- Tong, N., Lu, H., Ruan, X., & Yang, M. H. (2015). Salient object detection via bootstrap learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1884-1892).
- Qiu, W., Gao, X., & Han, B. (2017). A superpixel-based CRF saliency detection approach. *Neurocomputing*, 244, 19-32.
- Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International journal of computer vision*, 42(3), 145-175.
- Peng, H., Li, B., Ling, H., Hu, W., Xiong, W., & Maybank, S. J. (2017). Salient object detection via structured matrix decomposition. *IEEE transactions on pattern analysis and machine intelligence*, 39(4), 818-832.
- Lafferty, J., McCallum, A., & Pereira, F. C. (2001). Conditional random fields: Probabilistic models for segmenting and labeling sequence data.
- Jiang, H., Wang, J., Yuan, Z., Wu, Y., Zheng, N., & Li, S. (2013). Salient object detection: A discriminative regional feature integration approach. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2083-2090).

- Fu, K., Gong, C., Yang, J., Zhou, Y., & Gu, I. Y. H. (2013). Superpixel based color contrast and color distribution driven salient object detection. *Signal Processing: Image Communication*, 28(10), 1448-1463.
- Margolin, R., Tal, A., & Zelnik-Manor, L. (2013). What makes a patch distinct?. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1139-1146).
- Jiang, H., Wang, J., Yuan, Z., Liu, T., Zheng, N., & Li, S. (2011, September). Automatic salient object segmentation based on context and shape prior. In *BMVC* (Vol. 6, No. 7, p. 9).
- Zhu, W., Liang, S., Wei, Y., & Sun, J. (2014). Saliency optimization from robust background detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2814-2821).
- Li, Y., Hou, X., Koch, C., Rehg, J. M., & Yuille, A. L. (2014). The secrets of salient object segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 280-287).
- Sun, J., Lu, H., & Liu, X. (2015). Saliency region detection based on Markov absorption probabilities. *IEEE Transactions on Image Processing*, 24(5), 1639-1649.
- Wang, L., Lu, H., Ruan, X., & Yang, M. H. (2015). Deep networks for saliency detection via local estimation and global search. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3183-3192).
- Zhao, R., Ouyang, W., Li, H., & Wang, X. (2015). Saliency detection by multi-context deep learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1265-1274).

استناد به این مقاله:

شناسه دیجیتال (DOI): 10.22091/jemsc.2018.1275

شوریایی، محمد، فدائی اسلام، محمد جواد. (۱۳۹۷). «بهبود الگوریتم تشخیص نقشه برجستگی مبتنی بر CRF با استفاده از ویژگی‌های مبتنی بر تجزیه ماتریس». *مدیریت مهندسی و رایانش نرم*، ۶(۲)، ۱۶۶-۱۵۱.