

رویکردی مبتنی بر میدان تصادفی شرطی برای ادغام ویژگی‌ها در تشخیص برجستگی

محمد شوریابی و محمدجواد فدائی اسلام*

چکیده

توانایی تشخیص تمرکز انسان هنگام نگاه کردن به یک صحنه ثابت، یکی از موارد مورد علاقه جامعه بینایی ماشین و پردازش تصویر است. علیرغم اینکه تاکنون کارهای زیادی در این زمینه انجام شده است، اما با توجه به تنوع و پیچیدگی تصاویر، نیاز به تحقیق بیشتر در این زمینه کاملاً محسوس است. در این مقاله یک الگوریتم کارا و با ناظر برای شناسایی نقشه برجستگی با استفاده از میدان تصادفی شرطی ارایه شده است. تلفیق روش‌های تشخیص نقشه برجستگی پایه و روش‌های مبتنی بر تجزیه ماتریس به عنوان ورودی میدان تصادفی شرطی از نوآوری‌های این مقاله به شمار می‌رود. ارزیابی روش پیشنهادی روی دو پایگاه داده ECSSD و MSRA-10k با توجه به معیارهای ارزیابی، حاکی از کارایی مناسب روش پیشنهادی است. از منظر دیگر این مقاله وزن‌های محاسبه شده با استفاده از میدان تصادفی شرطی را به عنوان ابزاری برای ارزیابی روش‌های نقشه برجستگی به کار برده است که با استفاده از آن می‌توان به نوعی انتخاب ویژگی نمود. با توجه به با ناظر بودن میدان تصادفی شرطی تعداد ۲۰ تصویر که به تصویر ورودی نزدیک هستند برای آموزش میدان تصادفی در مقالات مرسوم به کار می‌روند. ارزیابی جامع روش‌های انتخاب تصویر مشابه را می‌توان از دستاوردهای دیگر این مقاله در نظر گرفت.

کلیدواژه‌ها

تشخیص برجستگی، میدان تصادفی شرطی، ابرپیکسل، تجزیه ماتریس

۱ مقدمه

محلی در نظر گرفته می‌شود و معمولاً از تضادهای بین اشیا و محیط پیرامونشان، از قبیل تفاوت رنگ، بافت، شکل و غیره، به وجود می‌آید. این مکانیزم، به طور ذاتی محرک اصلی سیستم بینایی را می‌سازد و عمدتاً در مرحله اولیه مواجهه دیداری با تصویر ورودی، توجه انسان را جذب می‌کند [۳]. فرآیندهای بصری سطح بالا سعی دارند به طور خودکار اهمیت مناطق مختلف تصویر را تشخیص دهند و تنها بر روی «جسم برجسته» ای که عمدتاً با هدف مدنظر مرتبط است، تمرکز یابند [۵].

طی دهه‌های اخیر، محققان پژوهش‌های زیادی پیرامون تعامل مکانیزم توجه بینایی و مدل‌های محاسباتی آن انجام داده‌اند. تحقیقات تشخیص برجستگی در بیش از سال‌های ۲۰۱۰، در مرحله پیشرفت‌های بنیادی قرار داشتند و توجه رشته‌های مختلف از جمله روانشناسی شناختی، علوم اعصاب و بینایی کامپیوتری را به خود جلب نمودند. از سال ۲۰۱۰، به منظور تشخیص برجستگی، مدل‌های پیشرفته‌تر و به ویژه مدل‌های مبتنی بر گراف

با افزایش استفاده از دستگاه‌های هوشمند و شبکه‌های اجتماعی، هر روزه غرق در حجم زیادی از داده‌های رسانه‌های دیجیتال می‌باشیم. با توجه به کمبود زمان، انتخاب پیام‌های مفیدتر از بین تمام داده‌های موجود به منظور پردازش بیشتر، امری ضروری محسوب می‌شود. در این راستا شناسایی نواحی با اهمیت تصویر (نواحی نقشه برجسته) یکی از ابزارهای مفید است که می‌توان به وسیله آن مفاهیم بصری مهم را استخراج نمود.

در مراجع [۲] تا [۴] برجستگی معمولاً به عنوان کنتراست

این مقاله در تیرماه سال ۹۸ دریافت، دومین بازنگری در مردادماه ۹۹ انجام و در شهریورماه همان سال پذیرفته شد.

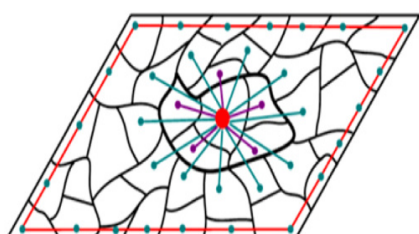
^۱ دانش‌آموخته کارشناسی ارشد، دانشکده برق و کامپیوتر، دانشگاه سمنان

رایانامه: m.shouryabi@semnan.ac.ir

^۲ دانشکده برق و کامپیوتر، دانشگاه سمنان، سمنان

رایانامه: fadaei@semnan.ac.ir

مرز تصویر به یکدیگر متصل هستند (خطوط قرمز در شکل (۱)). آن‌ها معتقدند بیشتر رویکردهای CRF موجود، مدل‌های گرافی احتمالی را با گراف پیکسل همسایگی ۸-گانه شبکه‌ای شکل تنظیم می‌کنند، که این سبب پیچیده شدن محاسبات می‌شود، در حالی که بررسی گراف سطح ابرپیکسل آن‌ها نه تنها می‌تواند مدل را ساده کند، بلکه باعث می‌شود هر ابرپیکسل، همسایگی و اطلاعات مرزی بیشتری داشته باشد که برای یادگیری و نتیجه‌گیری در مدل گراف مفید است.



شکل ۱ مدل گراف با ابرپیکسل‌ها به عنوان گره‌ها (خطوط بنفش و فیروزه‌ای به ترتیب اتصال به همسایگان اول و دوم را نشان می‌دهد. خط قرمز بدان معنی است که تمام گره‌های مرزی به هم متصل هستند.) [۱۱]

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 انتخاب می‌شود، سپس یک مدل میدان تصادفی شرطی خاص، برای این تصاویر که به تصویر ورودی نزدیک هستند، آموزش می‌بیند.

میدان تصادفی شرطی که Qiu و همکاران پیشنهاد می‌کنند، در واقع مدلی احتمالی برای محاسبه $P(A|F)$ می‌باشد که در آن مجموعه برجسته‌ها $F = \{f_1, f_2, f_3\}$ و ویژگی‌های برجستگی می‌باشد. در مقاله آن‌ها، CRF به صورت معادله (۱) تعریف می‌شود.

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

که در آن Z ثابت نرمال‌سازی به نام تابع افراز و $E(A|F)$ تابع انرژی است. آن‌ها سعی کردند، نشانه‌های برجستگی و سیستم همسایگی را در تابع انرژی مورد نظر ایجاد کنند. بنابراین تابع انرژی را به صورت معادله (۲) تعریف کردند.

$$E(A|F) = \sum_{x \in S} \sum_{i=1}^m \lambda_i V_i(a_x, f_i) + \sum_{x \in S, x' \in N_b(x)} N(a_x, a_{x'}, F) \quad (2)$$

که در آن V_i و N به ترتیب پتانسیل وابستگی و اثر متقابل (تعامل) نامیده می‌شوند. به طور ویژه، پتانسیل وابستگی در

مطرح شدند که به طور کلی دقت کلی تشخیص را بهبود می‌بخشند. اگر شروع بررسی‌ها در مورد برجستگی تصویر را سال ۱۹۸۰ در نظر بگیریم، هر چه به سمت جلو پیش می‌رویم، رویکردهای یادگیری با ناظر رواج بیشتری یافته است. در مقاله Yang و همکاران [۶] برای تمایز برجستگی بصری از یادگیری مشترک میدان تصادفی شرطی (CRF) و دیکشنری بصری استفاده شده است. Mai و همکاران [۷]، برای تجمع 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. Tong و همکاران [۸]، از چند نمونه از نقشه برجستگی ضعیف به عنوان مجموعه آموزشی ماشین‌های بردار پشتیبان (SVM) و متعاقباً از آن‌ها برای ایجاد یک نقشه برجسته قوی استفاده می‌کنند. در این مقاله سعی شده است تا یک مدل با ناظر با استفاده از میدان تصادفی شرطی برای تشخیص نواحی برجسته تصویر ارایه شود.

ساختار مقاله به صورت زیر ساماندهی شده است. از آنجا که پیشینه مربوط به روش‌های تشخیص برجستگی در مقاله [۱] به تفصیل آمده است، در بخش (۲) تنها تحقیقاتی که به طور کامل مرتبط هستند مورد بررسی قرار گرفته‌اند. در بخش (۳) روش پیشنهادی تلفیقی و نتایج پیاده سازی آن آمده است. بخش (۴) به میدان تصادفی شرطی از منظر یک روش انتخاب ویژگی نگریسته است و سعی نموده تا عوامل مهم در تعیین نقش برجستگی موثر را استخراج نماید. بخش (۵) روش gist که برای تعیین تصاویر مرتبط با تصویر ورودی به کار می‌رود را مورد بررسی و ارزیابی قرار داده است. جمع‌بندی فعالیت‌های انجام شده در بخش (۶) آمده است.

۲ تحقیقات مرتبط

مدل اصلی پیشنهادی در این مقاله در واقع از تلفیق دو روش ارایه شده در مقالات [۹] و [۱۰] است که در این بخش به توضیح آن‌ها می‌پردازیم.

۱-۲ رویکرد CRF مبتنی بر ابرپیکسل برای تشخیص برجستگی

Qiu و همکاران [۹] یک الگوریتم ساده و با ناظر را برای شناسایی نقشه برجستگی با استفاده از میدان‌های تصادفی شرطی و ویژگی‌های برجستگی (کنتراست محلی^۱، مرکز سوگیری^۲ و پس‌زمینه^۳) پیشنهاد کرده‌اند. در مدل پیشنهادی، یک گراف سطح ابرپیکسل ایجاد می‌شود. در گراف ایجاد شده، هر گره (ابرپیکسل)، علاوه بر اتصالش با گره‌های همسایه (همسایه اول)، به گره‌هایی که مرز مشترک با گره‌های همسایه اول دارند (خطوط فیروزه‌ای در شکل (۱))، متصل است. همچنین، همه گره‌ها در

^۱Local contrast

^۲Center bias

^۳Backgroundness

ساختار درختی را به همراه اطلاعات اولیه‌ای از میزان برجستگی نواحی به مساله اضافه می‌کند.

Peng و همکاران در این مقاله سعی دارند اطلاعات رتبه ماتریس، اطلاعات مربوط به همسایگی و اطلاعات اولیه مربوط به نقشه برجستگی را به نوعی با هم ترکیب نمایند. آن‌ها از رابطه (۴) برای انجام موارد بیان شده بهره می‌گیرند، در این مقاله، نقشه برجستگی تصویر (F) از تجزیه ماتریسی به L (ماتریس مرتبه پایین^۲) که نشان‌دهنده پس زمینه تصویر و S (ماتریس تنک^۳) که نشان‌دهنده شی برجسته می‌باشد، محاسبه می‌شود.

$$\min_{L,S} \psi(L) + \alpha\Omega(S) + \beta\Theta(L,S) \quad s.t. \quad F = L + S \quad (4)$$

در این رابطه ψ نشان‌دهنده رتبه ماتریس L است، تابع Ω دربردارنده 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.

۳ مدل پیشنهادی تلفیقی

میدان تصادفی شرطی که ما در روش پیشنهادی استفاده می‌کنیم، برگرفته از مقاله [۹] است. CRF پیشنهادی در واقع مدلی احتمالی برای محاسبه $P(A|F)$ می‌باشد که در آن $A = \{a_x\}$ مجموعه برجسب‌ها و $F = \{f_1, f_2, f_3, f_4\}$ ویژگی‌های برجستگی هستند. CRF پیشنهاد شده به صورت معادله (۱) تعریف می‌شود. تفاوت روش ما با روش پیشنهادی ارایه شده در مقاله [۹] به این است که آن‌ها از سه نشانه‌ی برجستگی کنتراست محلی، مرکز سوگیری و پس‌زمینه‌ای استفاده می‌کنند، در حالی که ما علاوه بر نشانه‌های برجستگی بیان شده، خروجی مدل Peng و همکاران (SMD) را به عنوان ویژگی چهارم به مدل مورد نظر می‌دهیم. در قسمت آموزش، روش پیشنهادی ما دقیقاً مانند روش [۹] عمل می‌کند و با توجه به ۲۰ تصویری که برای آموزش در نظر گرفته شده و موجود بودن مبنای درستی تصاویر آموزش و ویژگی‌های برجستگی هر تصویر، مقادیر λ ها برای هر نشانه برجستگی را مشخص می‌کند.

تفاوت دیگر کار ما با مقاله [۹] در بخش استنتاج است Qiu، و همکارانش مقدار a_x را به صورت تصادفی تولید می‌کنند. مشکل اینجاست که یک تصویر تقریباً به ۲۰۰ تا ۲۵۰ ابرپیکسل تقسیم می‌شود، بنابراین شرایط اولیه مساله شامل ۲۰۰ تا ۲۵۰ عدد تصادفی است که این اعداد تصادفی تاثیر مستقیم در رابطه ارائه شده دارد. در واقع این اعداد در قسمت اول رابطه (۲) در ویژگی‌های به دست آمده از تصویر ضرب می‌شود که صفر بودن این عدد تصادفی باعث نادیده گرفتن ویژگی‌های به دست آمده از تصویر می‌شود. بنابراین ایجاد شرایط اولیه تصادفی برای هر ناحیه

ناحیه x ، تابعی از تمام ویژگی مشاهده شده (نشانه برجستگی) F و همین‌طور برجسب A می‌باشد. در واقع این تابع تاثیر ویژگی‌های برجستگی را روی برجسب برجستگی (برجسته بودن یا نبودن) نشان می‌دهد. Qiu و همکاران پتانسیل وابستگی را به صورت $V_i(a_x, f_i) = f_{ix}a_x$ تشکیل می‌دهند که در آن f_{ix} مقدار ویژگی برجستگی i در ابرپیکسل x است و a_x مقداری باینری دارد که تعلق یک ابرپیکسل x به ناحیه برجسته یا زمینه را نشان می‌دهد. 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 Qiu و همکاران به صورت معادله (۳) تعریف می‌شود.

$$N(a_x, a_{x'}, F) = 1(a_x \neq a_{x'}) \exp(-\beta \|F_x - F_{x'}\|) \quad (3)$$

Qiu و همکاران برای قسمت آموزش CRF، تعداد ۲۰ تصویر را به عنوان کاندیداهایی که صحنه مشابهی با تصویر ورودی دارند، از کل دیتاست توسط الگوریتم gist به دست می‌آورند. سپس هر کدام از تصاویر آموزش را توسط الگوریتم SLIC ناحیه‌بندی می‌کنند. پس از ناحیه‌بندی توسط الگوریتم SLIC، سه ویژگی برجستگی کنتراست محلی، مرکز سوگیری و پس‌زمینه‌ای را برای هر ناحیه از تصاویر آموزش استخراج می‌کنند. در نهایت CRF را با توجه به ۲۰ تصویری که برای آموزش در نظر گرفته شده و موجود بودن مبنای درستی^۱ و نشانه‌های برجستگی هر تصویر، آموزش می‌دهند. در این بخش CRF آموزش می‌بیند، تا مقادیر پارامترها (λ ها) برای بخش استنتاج مشخص شود.

در بخش استنتاج، ابتدا تصویر ورودی را توسط الگوریتم SLIC ناحیه‌بندی می‌کنند. پس از ناحیه‌بندی تصویر، ویژگی‌های برجستگی برای هر ناحیه از تصویر ورودی را مشخص می‌کنند. باید به این نکته توجه داشت که در مرحله آموزش مقدار a_x هر تصویر مشخص است اما در مرحله استنتاج این مقدار به عنوان هدف مساله است که باید توسط میدان تصادفی شرطی تعیین گردد. Qiu و همکاران به عنوان شرایط اولیه مقدار a_x را به صورت تصادفی تولید می‌کنند. در نهایت CRF با توجه به ویژگی‌های تصویر ورودی، پارامترهای محاسبه شده در قسمت آموزش و شرایط اولیه، نقشه برجستگی تصویر ورودی را به عنوان خروجی، نتیجه می‌دهد.

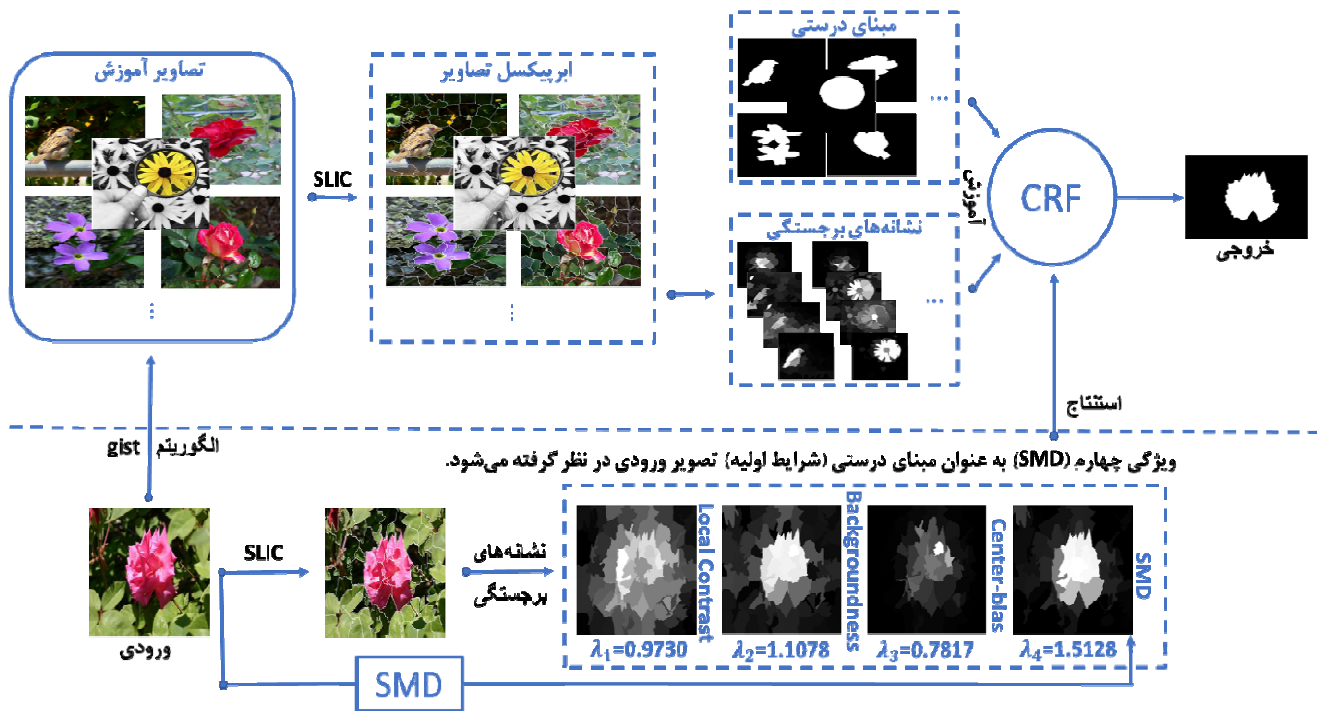
۲-۲ تشخیص شی برجسته از طریق تجزیه ماتریس ساختار یافته

Peng و همکاران [۱۰] یک روش کارا با استفاده از تجزیه ماتریس‌ها به منظور یافتن نواحی برجسته ابداع نموده‌اند. آن‌ها 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 $53 \times N$ (تعداد ابرپیکسل‌هاست) تشکیل می‌دهند. سپس

²Low rank

³Sparse matrix

¹Ground truth



شکل ۲. چارچوب مدل برجستگی پیشنهادی

مشق شده و نقشه مرجع را تعیین کند. در این بخش برای ارزیابی روش پیشنهادی از معیارهای میانگین خطای قدر مطلق^۱ (MAE)، نسبت همپوشانی^۲ (OR)، سطح زیر منحنی ROC^۳ (AUC) و F-measure استفاده می‌کنیم. WF نشان دهنده معیار F با β برابر با ۱ است و در WFb مقدار β برابر ۰٫۳ است. همانطور که در جدول (۱) مشاهده می‌شود. روش پیشنهادی ما در اغلب موارد نسبت به سایر روش‌ها بهتر عمل کرده است و این حاکی از کارایی مناسب آن نسبت به سایر روش‌ها است. شرح نحوه محاسبه و کارکرد معیارها در [۵] و [۹] آمده است.

جدول (۲) میانگین زمانی اجرای الگوریتم برای هر تصویر را در سیستمی با مشخصات 20.00GB (RAM) و Intel(R) Xeon(R) CPU E5640 @ 2.67GHz نشان می‌دهد. به علت در دسترس نبودن کد مقاله [۹] مقایسه زمانی با آن امکان پذیر نبود.

جدول ۲. میانگین زمان اجرای روش پیشنهادی	
روش‌ها	زمان (ثانیه)
روش پیشنهادی	۹٫۹۳
SMD	۳٫۲۸
PCA	۶٫۰۰

از تصویر، می‌تواند مدل را گمراه کند. پیشنهادی که برای این قسمت از الگوریتم داریم، این است که با توجه به اینکه CRF به هر یک از ویژگی‌های برجستگی موجود وزنی را اختصاص می‌دهد، ما ویژگی که وزن بیشتری دارد را به عنوان شرایط اولیه تصویر ورودی در نظر بگیریم. بنابراین در قسمت استنتاج، با توجه به وزن‌های محاسبه شده در قسمت آموزش، نشاندهای برجستگی موجود برای تصویر ورودی و همچنین مقدار شرایط اولیه که نحوه محاسبه آن ذکر شد، برجستگی تصویر ورودی محاسبه می‌شود (شکل ۲).

□ پیاده سازی و نتایج

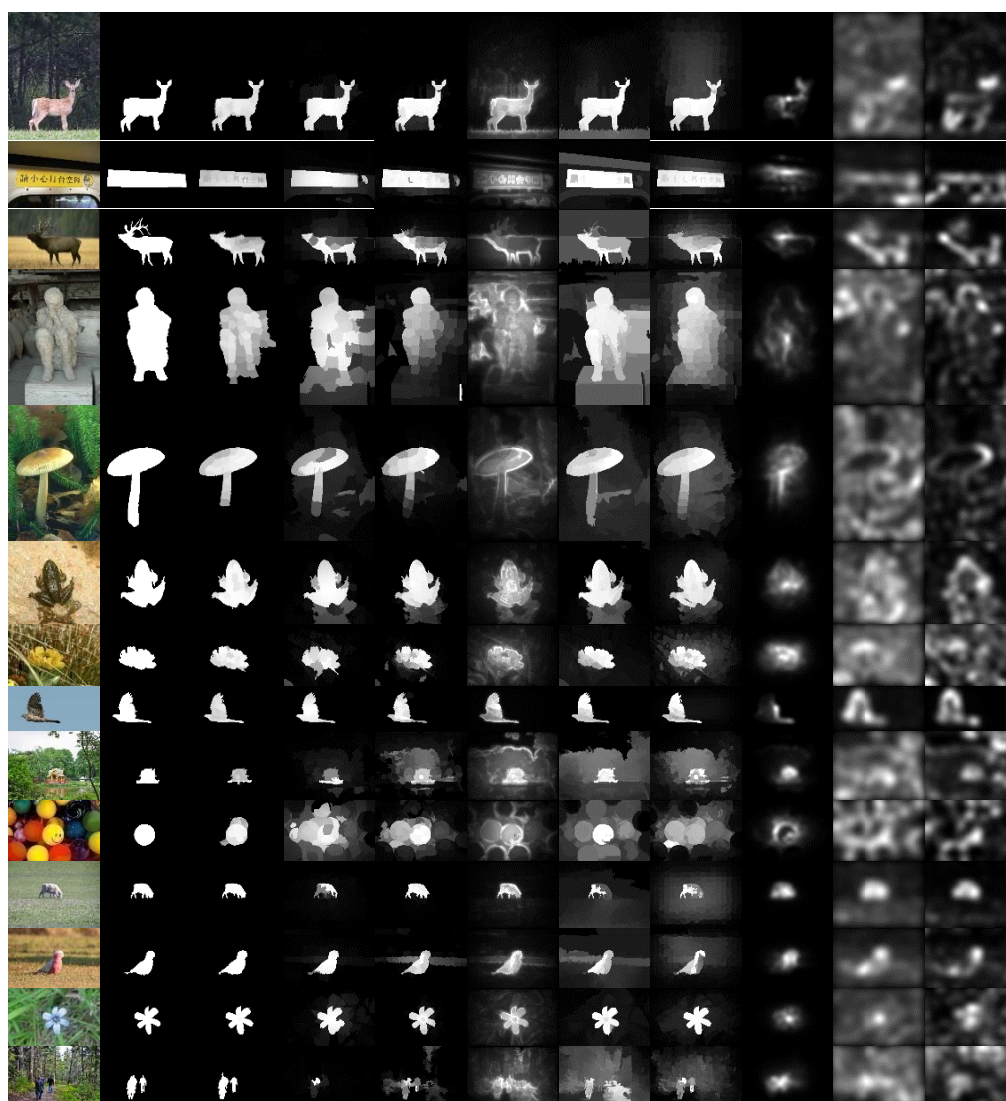
مدل پیشنهادی تلفیقی (الگوریتم ۱) که از ترکیب یک روش با ناظر با یک روش تجزیه ماتریس بدون ناظر ایجاد شده است، توسط نرم افزار MATLAB R2017a پیاده سازی و بر روی پایگاه داده MSRS-10K و ECSSD ارزیابی و با مدل‌های تشخیص برجستگی اخیر شامل مدل‌های SMD [۱۰]، RBD [۱۲]، PCA [۱۳]، SR [۱۴]، CHS [۱۵]، MC [۱۶]، COV [۱۷]، SS [۱۸]، SEG [۱۹] و HCT [۲۰] مقایسه شد. مقایسه بصری نقشه‌های برجستگی مدل پیشنهادی نسبت به مدل‌های دیگر در شکل (۳) نشان داده شده است. همانطور که در شکل مشاهده می‌کنید، نقشه برجستگی حاصل از مدل پیشنهادی، به مبنای درستی تصویر مورد نظر نزدیک‌تر است. 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

^۱Mean Absolute Error

^۲Overlap Ratio

^۳Area Under Curve



ورودی مبنای روش SMD RBD PCA CHS MC COV SS SR
درستی پیشنهادی

شکل ۳ مقایسه بصری نقشه‌های برجستگی مدل پیشنهادی با مدل‌های دیگر

الگوریتم ۱: مدل پیشنهادی تلفیقی

- (۱) خواندن تصویر ورودی (=تست)
- (۲) یافتن ۲۰ تصویر مشابه به تصویر ورودی (=تصاویر آموزش) % توسط الگوریتم gist
- (۳) بخش‌بندی تصاویر مربوط به آموزش و تصویر ورودی % توسط الگوریتم SLIC
- (۴) استخراج ویژگی‌ها local contrast, backgroundness, center-bias و خروجی مقاله پنگ و همکاران) از هر ناحیه تصاویر آموزش و تصویر تست
- (۵) آموزش CRF با تصاویر آموزش
- (۶) آزمایش CRF آموزش دیده با تصویر تست

جدول ۱ مقایسه روش پیشنهادی با سایر روش‌ها با توجه به معیارهای ارزیابی

		روش پیشنهادی	Qiu و همکاران	SMD	RBD	PCA	MC	CHS	SS
ECSSD	MAE	۰,۱۴۴۳	۰,۱۴۵۲	۰,۱۷۵۴	۰,۱۷۱۴	۰,۲۴۷۱	۰,۲۰۲۴	۰,۲۲۵۶	۰,۳۴۳۴
	OR	۰,۵۴۸۸	-	۰,۵۳۲۸	۰,۵۲۴۹	۰,۳۹۵۰	۰,۵۳۰۶	۰,۴۶۶۲	۰,۱۸۶۸
	AUC	۰,۷۶۴۶	-	۰,۷۹۰۴	۰,۷۸۰۹	۰,۷۹۱۴	۰,۸۱۶۰	۰,۸۰۶۶	۰,۶۸۷۳
	WF	۰,۵۹۸۰	-	۰,۵۲۳۱	۰,۵۱۲۱	۰,۳۶۴۳	۰,۴۵۴۴	۰,۴۶۱۲	۰,۲۵۶۹
	Wfb	۰,۶۴۲۱	۰,۵۸۱۵	۰,۵۲۳۸	۰,۵۱۲۱	۰,۳۵۳۸	۰,۴۴۱۵	۰,۴۳۳۱	۰,۲۳۳۵
MSRA10K		روش پیشنهادی	Qiu و همکاران	SMD	RBD	PCA	SEG	HCT	HS
	MAE	۰,۰۷۵۲	۰,۰۸۶۰	۰,۱۰۴۶	۰,۱۱۱۰	۰,۱۸۵۳	۰,۱۸۵۳	۰,۱۴۶۲	۰,۱۴۹۲
	OR	۰,۷۱۹۳	-	۰,۷۴۱۰	۰,۷۱۰۳	۰,۵۷۶۳	۰,۵۷۶۳	۰,۶۷۴۱	۰,۶۵۶۲
	AUC	۰,۷۹۰۹	-	۰,۸۴۷۱	۰,۸۳۲۷	۰,۸۳۹۲	۰,۸۳۸۸	۰,۸۴۵۲	۰,۸۳۳۱
	WF	۰,۷۸۶۰	-	۰,۷۰۴۱	۰,۶۸۰۹	۰,۴۷۳۰	۰,۴۷۳۰	۰,۵۸۲۱	۰,۶۰۴۲
	Wfb	۰,۸۱۹۶	۰,۷۰۹۵	۰,۷۰۵۰	۰,۷۱۷۷	۰,۴۸۳۰	۰,۴۸۲۹	۰,۶۰۳۳	۰,۶۱۰۲

صورت دلخواه تشکیل شده‌اند و دسته سوم، چهارم و پنجم ویژگی‌های دارای وزن بالا در محاسبات CRF هستند. با دقت در جدول (۳) به نظر می‌رسد ضعیف‌ترین ویژگی ماتریس مرتبه پایین باشد چون عدد آن منفی شده است اما با مشاهده ماتریس مرتبه پایین در شکل (۴) می‌توان این برداشت را انجام داد که این ویژگی به خوبی برجستگی را تشخیص می‌دهد اما به صورت برعکس این کار را انجام داده است. بنابراین دسته پنجم ویژگی‌ها به سه ویژگی با اندازه عددی بیشتر اختصاص داده شد.

جدول ۳ ضریب تاثیر ویژگی‌ها

ویژگی	
کنتراست محلی	۱,۲۰۶۲
پس‌زمینه‌ای	۱,۳۰۸۰
مرکز سوگیری	۰,۹۸۵۰
دانش اولیه زمینه	۱,۵۴۱۳
دانش اولیه رنگ	۰,۵۴۶۳
دانش اولیه موقعیت	۱,۰۸۱۴
ماتریس مرتبه پایین	-۱,۶۳۰۸
ماتریس پراکنده	۱,۷۶۷۴

جدول ۴ ترکیب مختلف ویژگی‌ها موجود

شماره	ویژگی‌ها
۱	کنتراست محلی، پس‌زمینه‌ای، دانش اولیه موقعیت و دانش اولیه زمینه
۲	کنتراست محلی، پس‌زمینه‌ای، ماتریس مرتبه پایین و ماتریس پراکنده
۳	ماتریس پراکنده، دانش اولیه زمینه، پس‌زمینه‌ای
۴	ماتریس پراکنده، دانش اولیه زمینه
۵	ماتریس پراکنده، (ماتریس مرتبه پایین-۱)، دانش اولیه زمینه

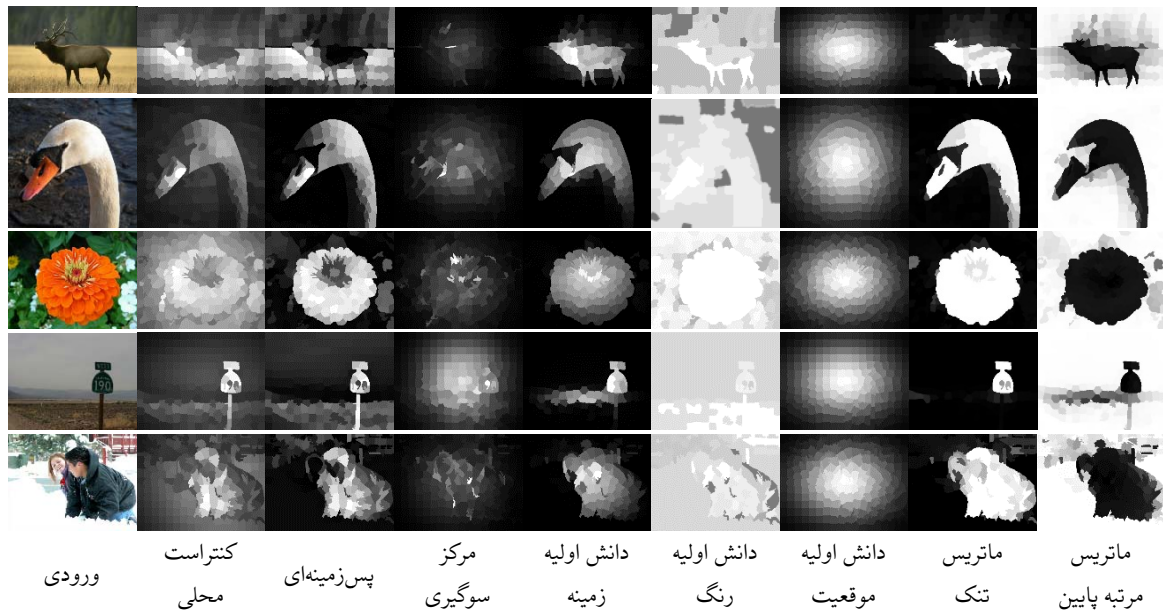
۴ رتبه‌بندی ویژگی‌ها با استفاده از CRF

در مرجع [۹] از سه روش پایه استخراج نقشه برجستگی کنتراست محلی، مرکز سوگیری و پس‌زمینه‌ای و در مقاله [۱۰] سه روش پایه دانش اولیه موقعیت^۱، دانش اولیه رنگ^۲ و دانش اولیه زمینه^۳ مورد استفاده قرار گرفته است. از سوی دیگر Peng و همکاران نقشه برجستگی تصویر را از تجزیه ماتریسی به L (ماتریس مرتبه پایین) که نشان‌دهنده پس زمینه تصویر و S (ماتریستک) که نشان‌دهنده شی برجسته می‌باشد، محاسبه می‌کنند [۱۰]. شکل (۴) این دو ماتریس را به صورت بصری و قابل درک نشان داده است.

استفاده از همه ویژگی‌ها باعث ایجاد بار محاسباتی بالا می‌شود، ترکیب مناسب از این ویژگی‌ها می‌تواند روند اجرای الگوریتم را سرعت بخشد. همانطور که قبلاً هم بیان شد، CRF به هر کدام از ویژگی‌های یک وزن نسبت می‌دهد که نشان دهنده میزان مناسب بودن آن است. بنابراین ما می‌توانیم چارچوب CRF پیشنهاد شده را برای ۸ ویژگی بیان شده (یعنی ویژگی‌های دو مقاله)، برای تعدادی تصویر آموزش دهیم و مقادیر λ (ضریب تاثیر) هر ویژگی را به دست آوریم. سپس ویژگی‌های که مقادیر λ بیشتر را دارند را برای سیستم پیشنهادی استفاده کنیم و سایر ویژگی‌ها را حذف کنیم. برای محاسبه λ ها، CRF برای ۲۰۰ تصویر آموزش داده شد، همانطور که در جدول (۳) مشاهده می‌شود، بهترین ویژگی از نظر CRF ماتریس پراکنده و بدترین ویژگی ماتریس مرتبه پایین است.

با توجه به ویژگی‌های موجود، پنج دسته ویژگی انتخاب شد که در جدول (۴) مشخصات آن آمده است. دو دسته اول به

¹Location prior²Color prior³Background prior



شکل ۴ مقایسه بصری ویژگی‌ها

۵ بررسی الگوریتم gist در روش پیشنهادی

شبکیه چشم انسان عمدتاً با دو نوع سلول بصری شناخته شده و پر می‌شوند: مخروطی و میله‌ای. 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. در نگاه اول به یک تصویر، سیستم بصری انسان تنها ساختار، توزیع رنگ و اطلاعات بافت را ثبت می‌کند. بنابراین با توجه به مواردی که بیان شد، Qiu و همکاران [۹] 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 و جو دارند را با استفاده از ویژگی gist پیدا می‌کنند.

به طور خاص، ویژگی gist به عنوان یک ویژگی سراسری برای بازیابی کاندیداها معرفی شده است. برای اینکه تاثیر استفاده از gist را در الگوریتم پیشنهادی مشخص کنیم. روش پیشنهادی تلفیقی را یک بار با استفاده از تصاویر مشابه به دست آمده توسط الگوریتم gist و یک بار با استفاده از تصاویری که از نظر gist کمترین تشابه را با تصویر ورودی دارند، آموزش می‌دهیم. در واقع در هر دو اجرا تعداد ۲۰ تصویر را برای آموزش CRF استفاده می‌کنیم، با این تفاوت که در اجرای اول ۲۰ تصویری که بیشترین تشابه را دارند (البته از نظر الگوریتم gist) و در اجرای دوم ۲۰ تصویری که کمترین تشابه را دارند، انتخاب می‌شوند. شرایط هر دو اجرا کاملاً یکسان است و این اجرا روی دیتاست ECSSD انجام شده است. اما انتظار می‌رود نتایج دو اجرا با هم اختلاف زیادی داشته باشند، اما همانطور که در جدول (۶) مشاهده می‌کنید، نتایج هر دو اجرا

پیاده سازی و نتایج

تمامی مواردی که در این بخش بیان شد، توسط نرم افزار MATLAB R2017a پیاده سازی و بر روی پایگاه داده ECSSD و MSRS-10K ارزیابی شده است. نتایج حاصل از هر ترکیب در جدول (۵) گزارش شده است. همانطور که مشاهده می‌شود، ترکیب شماره ۵ که با توجه به مقادیر λ به دست آمده توسط CRF پیشنهاد شد، از بقیه ترکیب‌ها بهتر عمل کرده است.

جدول ۵ نتایج حاصل از ترکیب ویژگی‌های مختلف

روشتلفیه ی	۱	۲	۳	۴	۵	
ECSSD	MA E	۰٫۱۶ ۲	۰٫۱۴ ۳	۰٫۱۳ ۴	۰٫۱۳ ۴	۰٫۱۴۴
	OR	۰٫۵۱ ۰	۰٫۵۲ ۷	۰٫۵۵ ۹	۰٫۵۵ ۸	۰٫۵۴۸
	AU C	۰٫۷۴ ۱	۰٫۷۴ ۴	۰٫۷۶ ۳	۰٫۷۵ ۹	۰٫۷۶۴
	WF	۰٫۵۹ ۰	۰٫۶۰ ۹	۰٫۶۴ ۳	۰٫۶۴ ۵	۰٫۵۹۸
	WFb	۰٫۵۸ ۰	۰٫۶۲ ۱	۰٫۶۴ ۵	۰٫۶۴ ۲	۰٫۶۴۲
MSRA10 K	MA E	۰٫۰۹ ۲	۰٫۰۸ ۰	۰٫۰۸ ۰	۰٫۰۸ ۱	۰٫۰۷۵
	OR	۰٫۶۷ ۱	۰٫۶۸ ۳	۰٫۷۰ ۰	۰٫۶۹ ۵	۰٫۷۱۹
	AU C	۰٫۸۰ ۴	۰٫۷۸ ۹	۰٫۸۰ ۱	۰٫۸۰ ۰	۰٫۷۹۰
	WF	۰٫۷۳ ۹	۰٫۷۷ ۸	۰٫۷۷ ۳	۰٫۷۶ ۸	۰٫۷۸۶
	WFb	۰٫۷۱ ۱	۰٫۷۹ ۹	۰٫۷۸ ۲	۰٫۷۷ ۴	۰٫۸۱۹

بیشترین فاصله را از تصویر ورودی دارند، آموزش داده شد. نتایج حاصل از این دو اجرا را در جدول (۸) مشاهده می‌کنید. نتایج همانطور که پیش بینی می‌شود متفاوت است، البته این تفاوت زیاد نیست. از همین تفاوت کم هم می‌توان نتیجه گرفت که ویژگی‌های انتخاب شده در کارایی gist تاثیر گذار هستند. اما باز هم آن تاثیری که ما انتظار داشتیم و مقاله Qiu و همکاران روی آن مانور داده بودند، برآورده نشد.

به طور خلاصه، اغلب ویژگی‌های که در این مقاله بررسی شده‌اند، تقریباً اثر مشابه دارند، یعنی به طور مثال یک ویژگی برای همه تصاویر خیلی خوب عمل می‌کند، یا به عبارتی دارای λ بزرگ‌تر است و دیگری برای همه ضعیف‌تر عمل می‌کند (ترتیب ضریب تاثیر آن‌ها حفظ می‌شود). انتخاب ویژگی‌های که حساسیت‌های متفاوتی برای صحنه‌های مختلف ایجاد نمایند تا اثر انتقال تجربه را بهتر نشان دهند، می‌تواند موضوع تحقیقات بعدی باشد.

جدول ۷ مقایسه λ ها روش تلفیقی و روش ترکیبی شماره ۳

روش تلفیقی	۰,۸۳۲	۰,۷۵۳	۰,۷۵۳	۱,۷۵۰
روش ترکیبی ۳	۱,۰۱۱	۱,۰۲۹	۱,۰۳۸	

جدول ۸ مقایسه روش ترکیبی (۳) با استفاده از ۲۰ تصویر با بیشترین/کمترین تشابه

ECSSD	توضیحات		۱	۲
	۱ - روش ترکیبی (۳) با استفاده از ۲۰ تصویر با بیشترین تشابه	MAE	۰,۱۳۴	۰,۱۴۹
		OR	۰,۵۵۹	۰,۵۳۷
		AUC	۰,۷۶۳	۰,۷۶۰
		WF	۰,۶۴۳	۰,۶۱۹
		WFb	۰,۶۴۵	۰,۶۲۷
	۲ - روش ترکیبی (۳) با استفاده از ۲۰ تصویر با کمترین تشابه			

۶ نتیجه‌گیری

در این مقاله روش پیشنهادی Qiu و همکاران در یافتن نواحی برجسته با استفاده از مدل پیشنهادی SMD بهبود داده شد. در این راستا، خروجی مقاله Peng را به عنوان ویژگی چهارم به چارچوب پیشنهادی Qiu و همکاران اضافه کردیم. البته علاوه بر اضافه کردن ویژگی جدید SMD، تغییراتی در نحوه محاسبه مبنای درستی ایجاد شد، که به طور مفصل نحوه محاسبه و دلیل این کار بیان شد. با توجه به ابداع ویژگی‌های متنوع در استخراج نواحی برجسته، روش CRF را به عنوان معیاری برای انتخاب ویژگی به کار بردیم که نتایج مفیدی به دست آمد. معمولاً از دیدگاه انتخاب ویژگی به روش CRF نگریده نمی‌شود که این موضوع می‌تواند از نوآوری‌های این مقاله باشد. در ادامه اثر الگوریتم gist به عنوان عاملی برای انتقال تجربه استخراج نواحی برجسته از تصاویر پردازش شده به تصویر هدف مورد بررسی قرار گرفت.

تقریباً یکسان است و به نظر می‌رسد که چه تصاویر مشابه انتخاب شوند و چه تصاویر که اصلاً تشابهی با تصویر ورودی ندارند، نتایج تغییر چندانی نمی‌کند. برای اطمینان بیشتر به جای gist، ۲۰ تصویر کاملاً تصادفی انتخاب شدند، باز هم در نتایج تغییری ایجاد نشد. بنابراین، تا اینجا با توجه به نتایج به دست آمده، می‌توان نتیجه گرفت که الگوریتم gist آن تاثیری را که از آن انتظار داشتیم و در [۹] بیان شده بود را نداشته است.

با توجه به نتایج به دست آمده می‌توان نظریات متفاوتی را مطرح کرد، از جمله اینکه gist، معیار مناسبی برای پیدا کردن تصاویر مشابه نیست. اما به نظر می‌رسد این نظر درست نیست، چون اگر مشکل از gist بود، نتایج حداقل برای ۲۰ تصویر تصادفی باید تغییر می‌کرد و بدتر یا بهتر می‌شود. اما نظر دیگری که می‌توان روی آن صحبت کرد، این است که gist زمانی می‌تواند کارایی خود را در انتقال تجربه نشان دهد که ویژگی‌های انتخاب شده جنبه‌های متفاوتی را پوشش دهند. در واقع آن چیزی که باعث شده نتایج تغییر نکنند، این است که برای ویژگی‌های موجود در روش تلفیقی تفاوتی نمی‌کند که ۲۰ تصویر انتخاب شده شبیه هستند یا نیستند، هر ۲۰ تصویری را که به CRF بدهیم، یک ویژگی، به عنوان مثال ویژگی SMD از بقیه بهتر عمل می‌کند. بنابراین زمانی gist می‌تواند خوب عمل کند که با عوض شدن تصاویر نزدیک به تصویر ورودی، مقدار λ به گونه‌ای تغییر کند که ویژگی بهتر، یک ویژگی غیر از آنچه قبلاً به دست آمده باشد.

جدول ۹ مقایسه روش پیشنهادی تلفیقی با استفاده از ۲۰ تصویر با بیشترین/کمترین تشابه از نظر Gist و تصاویر تصادفی

ECSSD	توضیحات		۱	۲	۳
	۱ - روش تلفیقی با استفاده از ۲۰ تصویر با بیشترین تشابه	MAE	۰,۱۴۷	۰,۱۴۹	۰,۱۴۹
		OR	۰,۵۴۸	۰,۵۴۷	۰,۵۳۶
		AUC	۰,۷۶۴	۰,۷۴۵	۰,۷۵۱
		WF	۰,۵۹۸	۰,۵۹۹	۰,۵۹۵
		WFb	۰,۶۴۲	۰,۶۳۷	۰,۶۳۱
	۲ - روش تلفیقی با استفاده از ۲۰ تصویر با کمترین تشابه				
	۳ - ۲۰ تصویر تصادفی				

برای اینکه مشخص شود آیا نظر بیان شده درست است یا خیر، به سراغ روش ترکیبی شماره ۳ (جدول (۴)) که شامل سه ویژگی ماتریس تنک، دانش اولیه زمینه و پس زمینه‌ای است، رفتیم. این دو روش (یعنی روش تلفیقی و روش ترکیبی) کاملاً از نظر ضریب تاثیر ویژگی‌ها با هم متفاوت هستند (جدول (۷)). در روش تلفیقی ویژگی چهارم که SMD است، اکثر مواقع بهتر عمل کرده و ضریب آن بیشتر شده است. بنابراین هر ۲۰ تصویری که به عنوان آموزش به CRF بدهیم، ویژگی SMD از بقیه بهتر عمل می‌کند. اما در روش ترکیبی شماره (۳) ویژگی‌ها تقریباً نزدیک به هم هستند و با عوض شدن تصاویر آموزش امکان تغییر نیز بیشتر می‌شود. مطابق آنچه برای روش پیشنهادی تلفیقی انجام شد، برای روش پیشنهادی ترکیبی نیز انجام شد و مدل مورد نظر یک بار با ۲۰ تصویر نزدیک به تصویر ورودی و یک بار با ۲۰ تصویر که

مراجع

- [15] J. Shi, Q. Yan, L. Xu and J. Jia, "Hierarchical Image Saliency Detection on Extended CSSD," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38, no. 4, pp. 717-729, 2015.
- [16] B. Jiang, L. Zhang, H. Lu, C. Yang and M.-H. Yang, "Saliency detection via absorbing markov chain," in Computer Vision, Sydney, 2013.
- [17] E. Erdem and A. Erdem, "Visual saliency estimation by nonlinearly integrating features using region covariances," J. Vision, vol. 13, 2013.
- [18] X. Hou, J. Harel and C. Koch, "Image Signature: Highlighting Sparse Salient Regions," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34, no. 1, pp. 194-201, 2011.
- [19] E. Rahtu, J. Kannala, M. Salo and J. Heikkilä, "Segmenting salient objects from images and videos," ECCV 2010: Computer Vision – ECCV, pp. 366-379, 2010.
- [20] J. Kim, D. Han, Y.-W. Tai and J. Kim, "Salient region detection via high-dimensional color transform," in IEEE Conference on Computer Vision and Pattern Recognition, Columbus, 2014.
- [1] شوریایی، م. فدائی اسلام، م. ج. "بهبود الگوریتم تشخیص نقشه برجستگی مبتنی بر CRF با استفاده از ویژگی‌های مبتنی بر تجزیه ماتریس،" مدیریت مهندسی و رایانش نرم، ۴ (۲) ۹۵-۱۱۰، ۱۳۹۷.
- [2] L. Itti, C. Koch and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," Pattern Analysis and Machine Intelligence, vol. 20, no. 11, pp. 1254 - 1259, 1998.
- [3] L. Itti and C. Koch, "Computational modelling of visual attention," Nature Reviews Neuroscience, vol. 2, no. 3, p. 194-203, 2001.
- [4] Y.-F. Ma and H. Zhang, "Contrast-based image attention analysis by using fuzzy growing," ACM Multimedia (ACMMM), pp. 374-381, 2003.
- [5] A. Borji, M.-M. Cheng, H. Jiang and J. Li, "Salient object detection: A benchmark," IEEE Transactions on Image Processing, vol. 24, no. 12, pp. 5706 - 5722, 2015.
- [6] J. Yang and M. Yang, "Top-Down Visual Saliency via Joint CRF and Dictionary Learning," IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 576-588, 2016.
- [7] L. Mai, Y. Niu and F. Liu, "Saliency aggregation: A data-driven approach," in Computer Vision and Pattern Recognition, Portland, 2013.
- [8] N. Tong, H. Lu, X. Ruan and M.-H. Yang, "Salient object detection via bootstrap learning," in Computer Vision and Pattern Recognition (CVPR), Boston, 2015.
- [9] W. Qiu, X. Gao and B. Han, "A superpixel-based CRF saliency detection approach," Neurocomputing, vol. 244, no. 28, pp. 19-32, 2017.
- [10] H. Peng, B. Li, H. Ling, W. Hu, W. Xiong and S. J. Maybank, "Salient Object Detection via Structured Matrix Decomposition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 4, pp. 818-832, 2016.
- [11] L. Zhang et al., "Ranking Saliency," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 9, pp. 1892 - 1904, 2017.
- [12] W. Zhu, S. Liang, Y. Wei and J. Sun, "Saliency Optimization from Robust Background Detection," in Computer Vision and Pattern Recognition, Columbus, 2014.
- [13] R. Margolin, A. Tal and L. Zelnik-Manor, "What Makes a Patch Distinct?," in Computer Vision and Pattern Recognition, Portland, 2013.
- [14] X. Hou and L. Zhang, "Saliency detection: A spectral residual approach," in Computer Vision and Pattern Recognition, Minneapolis, 2007.



محمد شوریایی دوره کارشناسی خود را در رشته مهندسی کامپیوتر- نرم افزار در دانشگاه تربت حیدریه گذراند. با پذیرش از طریق سهمیه استعداد درخشان وارد مقطع کارشناسی ارشد دانشگاه سمنان شد و در سال ۱۳۹۷ در رشته مهندسی کامپیوتر- هوش مصنوعی دانش آموخته گردید. زمینه تحقیقاتی ایشان یادگیری ماشین، شناسایی الگو، پردازش تصویر، شبکه‌های عصبی و منطق فازی است.



محمدجواد فدائی اسلام استادیار گروه مهندسی کامپیوتر دانشگاه سمنان است. ایشان دوره کارشناسی ارشد و دکترای مهندسی کامپیوتر- هوش مصنوعی را در دانشگاه علم و صنعت ایران گذرانده است. زمینه‌های مورد علاقه ایشان پردازش تصویر و ویدئو، شناسایی الگو و متن کاوی است.