

ردگیری اشیاء با استفاده از طبقه‌بند تک کلاسه و اطلاعات زمینه

سیده سعیده غفاریان طباطبایی^۱، امیرحسین طاهری نیا^۲، هادی صدوقی یزدی^۳، احسان شمس داوودلی^۴

^۱دانشکده مهندسی - دانشگاه فردوسی مشهد، saeede.ghaffarian@stu.um.ac.ir

^۲دانشکده مهندسی - دانشگاه فردوسی مشهد، taheerinia@yahoo.com

^۳دانشکده مهندسی - دانشگاه فردوسی مشهد، h-sadoghi@um.ac.ir

^۴دانشکده مهندسی - دانشگاه فردوسی مشهد، e_shams_d@yahoo.com

چکیده

امروزه ردگیری به‌عنوان مسأله طبقه‌بندی دودویی در نظر گرفته می‌شود که شیء را از پس‌زمینه متمایز می‌کند. برخی از طبقه‌بندهای دودویی که بدین منظور استفاده شده، توانایی اداره کردن نقاط پرت را ندارند که باعث کاهش عملکرد سیستم ردگیر شده است. طبقه‌بندهای تک‌کلاسه نیز به دلیل آن‌که تنها از نمونه‌های کلاس هدف استفاده می‌کنند، موجب مشکل بی‌مقصد رفتن هدف ردگیری شده می‌شود. در این مقاله از طبقه‌بند تک‌کلاسه همراه با نمونه‌های منفی استفاده می‌شود که اطلاعات صحنه در ساخت مدل هدف دخالت می‌کند. با توجه به نتایج و آزمایش‌ها، مشاهده می‌شود که استفاده از اطلاعات زمینه به‌عنوان نمونه‌های منفی در مقایسه با روشی که فقط از نمونه‌های مثبت استفاده می‌کند، باعث افزایش قابل ملاحظه کارایی ردگیر می‌شود.

واژه‌های کلیدی

ردگیری، طبقه‌بند تک‌کلاسه، اطلاعات زمینه.

۱- مقدمه

طبقه‌بندی دودویی نیمه ناظر برخط^۵ پیشنهاد شده‌اند. که در این روش‌ها علاوه بر استفاده از نمونه‌های برجسب‌دار، از نمونه‌های بدون برجسب نیز استفاده شده است تا توانایی خوبی برای اداره کردن نقاط پرت^۶ داشته باشد. با این وجود، جمع‌آوری داده‌های بدون برجسب باعث هزینه محاسبات بالا می‌شود [۳] از این‌رو عملکرد سیستم ردگیری در پردازش بلادرنگ^۷ کاهش می‌یابد. علاوه بر این، روش‌های ردگیری نیمه ناظر مشکل داده پرت را به‌طور کامل حل نمی‌کنند که منجر به مشکل بی‌مقصد رفتن^۸ هدف ردگیری شده می‌شود. در روش ارائه شده در [۴] برای حل این مشکلات، از طبقه‌بند تک‌کلاسه^۹ استفاده شده است که توانایی اداره کردن نقاط پرت و همچنین کنترل داده با ابعاد بالا را دارد. با این وجود، به دلیل در نظر گرفتن تنها نمونه‌های مثبت در طبقه‌بند SVDD همچنان مشکل بی‌مقصد رفتن

ردگیری^۱ مسأله‌ای مهم در زمینه بینایی کامپیوتر است که در چند دهه اخیر به‌طور گسترده مورد مطالعه قرار گرفته است. در حال حاضر، روش‌های ردگیری با تشخیص^۲، ردگیری را به‌صورت مسأله طبقه‌بندی دودویی در نظر می‌گیرند که در این حالت شیء از پس‌زمینه^۳ متمایز می‌شود. به عبارت دیگر، نواحی هدف به‌عنوان نمونه‌های مثبت و نواحی غیر هدف به‌عنوان نمونه‌های منفی در نظر گرفته می‌شوند و طبقه‌بند آموزش‌دیده مرز تصمیمی را پیدا می‌کند که مثبت‌ها و منفی‌ها را به بهترین نحو جدا نماید. از جمله روش‌های مورد استفاده برای طبقه‌بندی که توانایی خوبی برای کنترل داده در ابعاد بالا دارند، ماشین‌های بردار پشتیبان^۴ [۱] و یا طبقه‌بندهای به‌دست آمده توسط الگوریتم Adaboost [۲] هستند. همچنین روش‌های یادگیری نیمه ناظر نیز برای ردگیری شیء به‌عنوان یک مسأله

⁵ Online

⁶ Outlier

⁷ Real time

⁸ Drifting

⁹ SVDD

¹ Tracking

² Tracking-by-Detection

³ Background

⁴ Support Vector Machines

تنها نمونه‌های مثبت در نظر گرفته می‌شوند، که همچنان شاهد ضعف ردگیر در محیط‌های پیچیده هستند.

۳- روش پیشنهادی

همان‌طور که گفته شد، در مسأله ردگیری ویدئو، برای ساخت مدل هدف یک طبقه‌بند تک‌کلاسه در نظر گرفته شده است. در این روش نمونه‌های کلاس غیر هدف که به‌سادگی قابل دسترس هستند در ساخت مدل هدف تأثیر ندارند و برگرفته از مفهوم یادگیری زمینه، اطلاعات صحنه در یادگیری مدل هدف دخالت ندارد که این امر موجب بی‌مقصد رفتن هدف در ردگیری می‌شود. در روش ارائه شده، برای ساخت مدل هدف از طبقه‌بند تک‌کلاسه همراه با نمونه‌های منفی استفاده شده است.

۳-۱- طبقه‌بند تک‌کلاسه

فرض می‌شود $x_i (i = 1, \dots, n)$ نمونه‌های آموزشی p بعدی متعلق به یک کلاس هستند. ابرکره ای کمینه به مرکز $e = (e_1, e_2, \dots)^T$ و شعاع R در فضای ویژگی با بعد بالا شامل نقاط پرت به‌عنوان ناحیه کلاس در نظر گرفته می‌شود. بنابراین مسأله به صورت زیر فرمول‌بندی می‌شود [۷]:

$$\min R^2 + C \sum_{i=1}^n \xi_i \quad (1)$$

$$s.t. \|\varphi(x_i) - e\| \leq R^2 + \xi_i, \xi_i \geq 0, i = 1, 2, \dots, n$$

$\varphi(x)$ تابع نگاشتی است که x را به فضای ویژگی با بعد بالا نگاشت می‌دهد. $\xi = (\xi_1, \xi_2, \dots, \xi_n)^T$ که ξ_i متغیر کمبود نمونه آموزشی نام است و C قیدی است که بین اندازه ابرکره و خطای آموزش مصالحه ایجاد می‌کند.

بر این اساس در زمینه ردگیری نیز می‌خواهیم ابرکره‌ای پیدا کنیم که شامل بیشترین نمونه آموزش به‌دست‌آمده از ناحیه هدف است. بعد از مرحله آموزش، مرز تصمیم اجازه انتخاب ناحیه کاندید مناسب را می‌دهد.

۳-۲- طبقه‌بند تک‌کلاسه با نمونه‌های منفی

در این قسمت مدلی ارائه می‌شود که در فرآیند یادگیری علاوه بر نمونه‌های کلاس هدف (نمونه مثبت)، نمونه‌های یادگیری پرت (نمونه منفی) نیز دخالت داده می‌شوند [۸]. در این مدل سعی بر این است که کره‌ای با مرکز e و حداقل شعاع R یافت شود که نمونه‌های مثبت را در داخل خود قرار داده و از پذیرش نمونه‌های منفی خودداری کند. متغیرهای e و R مرتبط با طبقه‌بند NSVDD، جواب‌های مسأله بهینه‌سازی زیر هستند:

$$\min R^2 + C_1 \sum_{i=1}^n \xi_i^+ + C_2 \sum_{l=1}^m \xi_l^- \quad (2)$$

$$s.t. \|\varphi(x_i) - e\| \leq R^2 + \xi_i^+, \xi_i^+ \geq 0, i = 1, 2, \dots, n$$

$$\|\varphi(x_l) - e\| \leq R^2 - \xi_l^-, \xi_l^- \geq 0, l = 1, 2, \dots, m$$

در این مسأله n و m به ترتیب بیانگر اندیس نمونه‌های مثبت، اندیس نمونه‌های منفی، تعداد نمونه‌های مثبت و تعداد نمونه‌های منفی هستند. مطابق آنچه در قیود این مسأله مشاهده می‌شود سعی بر این است

هدف در ردگیری مشاهده می‌شود. ایده‌ای که به ذهن می‌رسد استفاده از اطلاعات زمینه^{۱۰} صحنه به‌عنوان نمونه‌های منفی است که باعث ارائه مدل بهتری از هدف می‌شود. در روش ارائه شده، از طبقه‌بند تک‌کلاسه همراه با نمونه‌های منفی^{۱۱} استفاده شده است که مدل پس‌زمینه را در ساختن مدل پیش‌زمینه^{۱۲} دخالت می‌دهد و مدل مقاوم‌تری از هدف به‌دست می‌آید. این روش مدلی متفاوت با طبقه‌بندهای دوکلاسه (مانند SVM) ارائه می‌دهد که توانایی اداره نقاط پرت را دارد همچنین با در نظر گرفتن نمونه‌های هر دو کلاس، یادگیری مقاوم‌تری نسبت به طبقه‌بند تک‌کلاسه ایجاد می‌شود.

ادامه مقاله به شرح زیر است: بخش ۲ مروری بر کارهای گذشته است. در بخش ۳ روش پیشنهادی مطرح شده و در بخش ۴ نتایج و آزمایش‌ها قرار داده شده است و در نهایت بخش ۵ نتیجه‌گیری ارائه شده است.

۲- مروری بر کارهای گذشته

روش‌های ردگیری با تشخیص خانواده‌ای از روش‌های یادگیری با ناظر یا نیمه ناظر می‌باشند که ردگیری را به‌صورت مسأله طبقه‌بندی دودویی در نظر می‌گیرند. در این ردگیری از ماشین بردار پشتیبان به‌عنوان طبقه‌بند استفاده می‌شود که دارای عملکرد مناسبی برای ردگیری وسایل نقلیه در حال حرکت است [۱]. باین‌حال در این روش SVM به‌صورت برخط به‌روزرسانی نمی‌شود که منجر به کاهش کارایی در ردگیری شده است. Bischof و Grabner از الگوریتم Adaboost [۲] برای اجرای برخط استفاده کردند که نمونه‌های مثبت و منفی را به‌صورت برخط جمع‌آوری می‌کنند. باین‌حال این روش‌ها ممکن است با وجود پس‌زمینه پیچیده دچار مشکل بی‌مقصد رفتن هدف در ردگیری شوند. Babenko و همکارانش [۵] از یادگیری چندگانه نمونه^{۱۳} استفاده کردند که منجر به نتایج خوبی در ردگیری شده است. Tang و همکارانش [۶] برای بهره‌گیری از ویژگی‌های مستقل چندگانه برای آموزش طبقه‌بندهای برخط، از co-training استفاده کردند. این طبقه‌بندها به‌طور مشترک داده‌های بدون برچسب را طبقه‌بندی می‌کنند و داده‌های برچسب‌دار جدید را برای به‌روزرسانی یکدیگر استفاده می‌کنند. هر ویژگی برای آموزش SVM برخط استفاده می‌شود و خروجی‌هایشان برای گرفتن نتایج نهایی طبقه‌بندی ترکیب می‌شوند. روش‌های ردگیری با تشخیص که در بالا ذکر شد، ردگیری را به‌عنوان مسأله طبقه‌بندی دودویی در نظر گرفته‌اند. مشکل رایج آن‌ها این است که نمی‌توانند به‌طور کامل نقاط پرت را اداره کنند که منجر به ردگیری نادرست می‌شود. همچنین یادگیری نیمه‌ناظر نیاز به تعداد زیادی نمونه بدون برچسب برای یادگیری دارد که باعث هزینه اضافی در استخراج ویژگی و طبقه‌بندی می‌شود. برای حل این مشکلات، Keren Fu و همکارانش [۴] از طبقه‌بند تک‌کلاسه استفاده کردند که در این طبقه‌بند

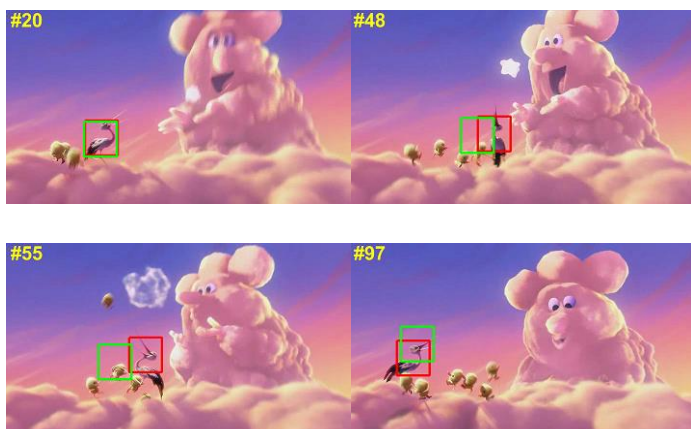
¹⁰ Contextual Information

¹¹ NSVDD

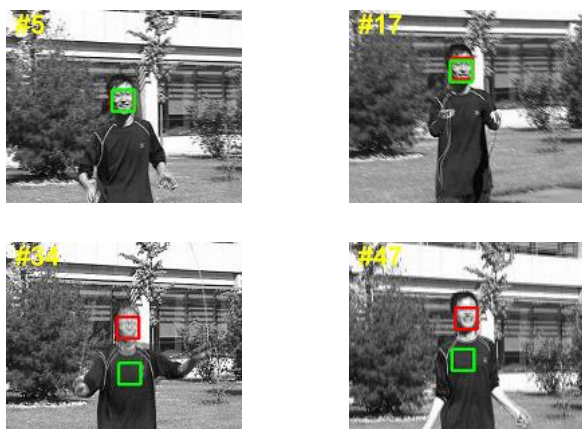
¹² Foreground

¹³ Multiple Instance Learning

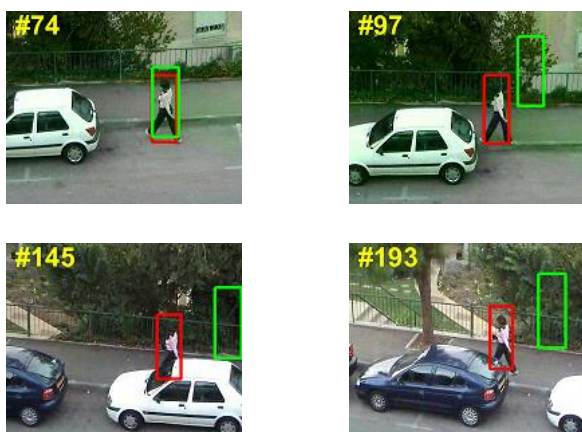
در این بخش نتایج به دست آمده از روش ارائه شده توسط این مقاله و روش [۴] که از SVDD استفاده کرده، روی چند مجموعه داده بررسی شده است. در شکل‌های زیر، مستطیل سبز ردگیر استفاده شده از مقاله [۴] و مستطیل قرمز، ردگیر ارائه شده در این مقاله است.



شکل ۱: نتیجه ردگیری بر روی مجموعه داده bird2.



شکل ۲: نتیجه ردگیری بر روی مجموعه داده jumping.



شکل ۳: نتیجه ردگیری بر روی مجموعه داده woman.

که ضمن پوشش نمونه‌های یادگیری مثبت، نمونه‌های یادگیری منفی از ابرکره بیرون رانده شوند. لیکن با در نظر گرفتن متغیرهای کمبود ξ_1^+ و ξ_1^- در مقابل رسیدن به مرزهای مناسب‌تر اجازه وقوع خطای یادگیری صادر می‌شود و پارامترهای C_1 و C_2 وظیفه تنظیم خطای یادگیری را بر عهده‌دارند.

در زمینه ردگیری نیز ابرکره‌ای پیدا می‌کنیم که شامل نمونه‌های کلاس هدف باشد در حالی که اطلاعات زمینه را در خود جای ندهد.

۳-۳- استخراج ویژگی

انتخاب ویژگی برای ردگیری از اهمیت بالایی برخوردار است. ویژگی‌های خوب معمولاً توانایی خوبی برای توصیف ظاهر هدف ردگیری شده دارند و در عین حال آن را از پس‌زمینه پیچیده و دیگر اشیاء جدا می‌کنند. در تحقیقات اخیر، بسیاری از انواع ویژگی‌ها، نظیر ویژگی‌های HOG^{۱۴} [۹]، هیستوگرام رنگ^{۱۵} [۱۰]، ویژگی‌های Haar-like [۱۱]، ویژگی‌های Gabor [۱۲] و ویژگی‌های LBP [۱۳] برای ردگیری اتخاذ شده‌اند. استفاده از هیستوگرام رنگ به‌عنوان انتخاب ویژگی در ردگیری، موجب بی‌هدف رفتن ردگیر می‌شود به دلیل اینکه هیستوگرام رنگ توانایی نسبتاً ضعیفی برای تمیز دادن هدف از پس‌زمینه دارد، مخصوصاً وقتی رنگ پس‌زمینه شبیه رنگ هدف باشد. به همین دلیل در این مقاله از هیستوگرام رنگ استفاده نشده و نوعی از ویژگی‌هایی که شکل و بافت هدف را توصیف می‌کنند انتخاب شده که نسبت به تغییرات روشنایی ثابت هستند. بنابراین در روش ارائه شده از استخراج ویژگی بر پایه HOG استفاده شده است.

۳-۴- چارچوب ردگیر

ردگیر پایه مبتنی بر فیلتر ذرات [۱۴] است. برای مدل حرکت در فیلتر ذرات از تبدیل آفین^{۱۶} استفاده شده که تغییر مقیاس و دوران را در خودش دارد و فضای حالت ذرات فیلتر به اندازه تعداد پارامترهای تبدیل آفین است. ردگیر، موقعیت هدف را در هر فریم به دست می‌آورد. در مرحله بعد کاندیدهای هدف اطراف موقعیت هدف ردگیر شده از فریم قبل با استفاده از فیلتر ذرات تولید می‌شوند. سپس وزن‌دهی به این کاندیدها با استفاده از آزمایش طبقه‌بند آموزش دیده صورت می‌گیرد. به روزرسانی طبقه‌بند نیز در هر فریم انجام می‌شود.

پارامترهای معادله ۲ در این مرحله به این صورت تنظیم می‌شود که C_1 نسبتاً بزرگ و C_2 بزرگ باشد تا به نمونه‌های مثبت اجازه خارج شدن و به نمونه‌های منفی اجازه وارد شدن به ابرکره داده نشود. پارامتر کرنل نیز به صورت مجزا برای هر ویدئو به دست می‌آید.

۴- نتایج آزمایش‌ها

¹⁴ Histogram of Oriented Gradients

¹⁵ Color Histogram

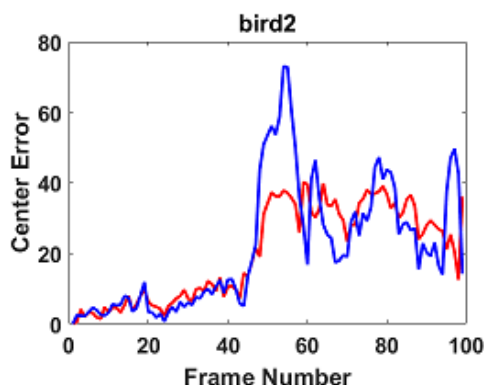
¹⁶ Affine

مثبت استفاده شده، کارایی ردگیری را به طور قابل ملاحظه‌ای افزایش می‌دهد.

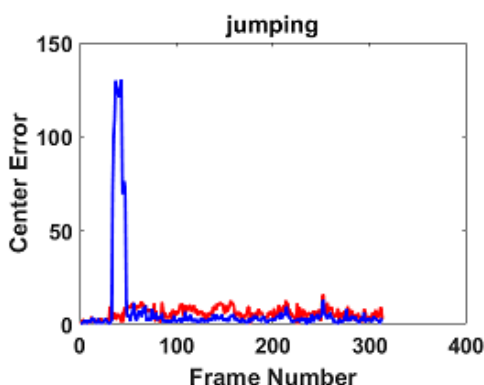
مراجع

- [1] S. Avidan, "Support vector tracking," *IEEE transactions on pattern analysis and machine intelligence*, vol. 26, pp. 1064-1072, 2004.
- [2] H. Grabner and H. Bischof, "On-line boosting and vision," in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, 2006, pp. 260-267.
- [3] X. Zhu and A. B. Goldberg, "Introduction to Semi-Supervised Learning (Synthesis Lectures on Artificial Intelligence and Machine Learning)," *Morgan and Claypool Publishers*, vol. 14, 2009.
- [4] K. Fu, C. Gong, Y. Qiao, J. Yang, and I. Y.-H. Gu, "One-class support vector machine-assisted robust tracking," *Journal of Electronic Imaging*, vol. 22, pp. 023002-023002, 2013.
- [5] B. Babenko, M.-H. Yang, and S. Belongie, "Visual tracking with online multiple instance learning," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, 2009, pp. 983-990.
- [6] F. Tang, S. Brennan, Q. Zhao, and H. Tao, "Co-tracking using semi-supervised support vector machines," in *2007 IEEE 11th International Conference on Computer Vision*, 2007, pp. 1-8.
- [7] Y. Forghani, S. Effati, H. S. Yazdi, and R. S. Tabrizi, "Support Vector Data Description by using hyper-ellipse instead of hyper-sphere," in *Computer and Knowledge Engineering (ICCKE), 2011 1st International eConference on*, 2011, pp. 22-27.
- [8] K. Lee, D.-W. Kim, D. Na, K. H. Lee, and D. Lee, "PLPD: reliable protein localization prediction from imbalanced and overlapped datasets," *Nucleic acids research*, vol. 34, pp. 4655-4666, 2006.
- [9] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 2005, pp. 886-893.
- [10] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean shift," in *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, 2000, pp. 142-149.
- [11] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, 2001, pp. I-511-I-518 vol. 1.
- [12] T. S. Lee, "Image representation using 2D Gabor wavelets," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 18, pp. 959-971, 1996.
- [13] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, pp. 971-987, 2002.
- [14] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on signal processing*, vol. 50, pp. 174-188, 2002.

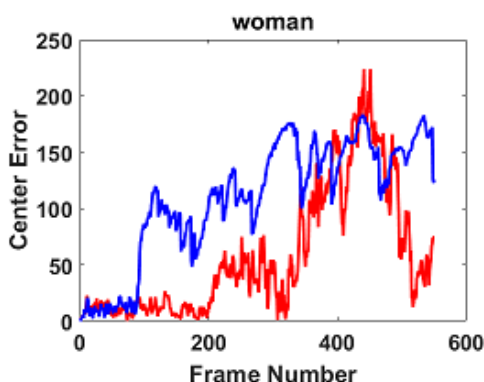
نمودارهای زیر خطای به‌دست‌آمده با این دو روش را در مقایسه با ground truth برای مجموعه داده‌های bird2، jumping و woman نشان می‌دهند که نمودار سبز خطای ردگیر استفاده‌شده از مقاله [۴] و نمودار قرمز، خطای ردگیر ارائه‌شده در این مقاله است.



شکل ۴: خطای ردگیری بر روی مجموعه داده bird2.



شکل ۵: خطای ردگیری بر روی مجموعه داده jumping.



شکل ۶: خطای ردگیری بر روی مجموعه داده woman.

۵- نتیجه‌گیری

در این مقاله، از طبقه‌بند تک‌کلاس همراه با اطلاعات زمینه برای ردگیری استفاده‌شده است. نتایج به‌دست آمده نشان می‌دهد استفاده از اطلاعات زمینه به‌عنوان نمونه‌های منفی در مقایسه با روشی که فقط از نمونه‌های