

روشی ترکیبی و نیمه نظارتی مبتنی بر گراف برای برچسب‌زنی خودکار تصاویر

مجتبی کردآبادی^۱، محرم منصوری زاده^۲ و حسن ختن‌لو^۳

۱ چکیده

روش‌های یادگیری نیمه نظارتی مبتنی بر گراف اغلب بر روی مسائل تک برچسبی متمرکز و پیاده‌سازی شده‌اند، در صورتی که بسیاری از مسائل دنیای واقعی به صورت چندبرچسبی هستند. در این مقاله یک روش نیمه نظارتی ترکیبی بنام LGC+ML-KNN برای برچسب‌زنی تصاویر به صورت چندبرچسبی ارائه داده‌ایم که از ترکیب روش یادگیری نیمه نظارتی مبتنی بر گراف (LGC) و یادگیری چندبرچسبی (ML-KNN) تشکیل شده است. روش ارائه‌شده به دلیل استفاده از یادگیری نیمه نظارتی و مشارکت دادن تمام نمونه‌ها و پیش‌بینی برچسب‌های اولیه و آموزش یادگیر ML-KNN با تعداد نمونه برچسب خورده بیشتر، دارای دقت بهتری نسبت به روش‌های موجود است. روش ارائه‌شده روی چندین مجموعه داده استاندارد آزمون شده است. نتایج آزمایش‌ها نشان می‌دهد که عملکرد روش ارائه‌شده بخصوص در مواردی که تعداد نمونه برچسب‌دار بسیار کم است، به نحو قابل ملاحظه‌ای بهتر است.

کلیدواژه‌ها: بازیابی تصاویر، برچسب‌زنی خودکار تصاویر، یادگیری نیمه نظارتی

(^۱TBIR) ۲- بازیابی بر اساس محتوا (CBIR)^۲. در پژوهش‌های حوزه CBIR سعی بر این است که با استخراج ویژگی‌های رنگ، بافت و شکل، تصاویری به کاربر نشان داده شود که بیشترین ارتباط را با تصویر پرس‌وجو داشته باشند. فهم و برداشتی که مدل کامپیوتری از تصویر دارد تنها بر مبنای ویژگی‌های استخراج‌شده است که عمل بازیابی نیز بر مبنای آن انجام می‌گیرد؛ بنابراین نتیجه بازیابی با نتیجه موردنظر انسان لزوماً مطابقت کامل نخواهد داشت. این مطلب که درک انسان و کامپیوتر از تصویر باهم تفاوت دارد به شکاف معنایی^۳ معروف است. در حال حاضر موتورهای جستجو در بازیابی داده‌های متنی دقت و سرعت مناسبی دارند. این امر باعث شده است تا پژوهشگران جهت بازیابی تصاویر، ابتدا مجموعه دادگان حاوی تصاویر را با اطلاعات متنی برچسب‌زنی کنند و سپس بازیابی را در حوزه متن انجام دهند. این کار سرعت عملیات بازیابی تصویر را تا حد قابل توجهی افزایش می‌دهد. برچسب‌زنی تصاویر به فرآیند تولید

۲ مقدمه

در فضای رو به رشد وب و شبکه‌های اجتماعی تعداد تصاویر روی وب نیز به‌گونه‌ای قابل توجه افزایش یافته است به طوری که سازمان‌دهی، جستجو و بازیابی تصاویر به یکی از چالش‌های اساسی در حوزه بازیابی اطلاعات تبدیل شده است. بازیابی تصاویر به دو روش انجام می‌گیرد: ۱- بازیابی بر اساس متن

این مقاله در دی‌ماه ۱۳۹۶ دریافت، در شهریورماه ۱۳۹۷ دومین بازنگری و در دی‌ماه همان سال پذیرفته شد.

^۱ دانشجوی کارشناسی ارشد، گروه مهندسی کامپیوتر، دانشگاه بوعلی سینا
رایانامه: kordabadi@basu.ac.ir
^۲ گروه مهندسی کامپیوتر، دانشگاه بوعلی سینا
رایانامه: mansoorm@basu.ac.ir
^۳ گروه مهندسی کامپیوتر، دانشگاه بوعلی سینا
رایانامه: hkh@basu.ac.ir

1 Text Based Image Retrieval
2 Content Based Image Retrieval
3 Semantic Gap

Archive of SID

بدون برچسب را برچسب‌زنی نماییم. با توجه به اینکه یکی از چالش‌های مهم در یادگیری نیمه نظارتی نحوه استفاده از دادگان بدون برچسب در فرآیند آموزش سیستم است، ما در مقاله حاضر توانسته‌ایم با استفاده از مدل یادگیری نیمه نظارتی مبتنی بر گراف دادگان بدون برچسب را در فرآیند آموزش دخالت دهیم و با دقت خوبی برچسب اولیه‌ای برای دادگان بدون برچسب پیش‌بینی کنیم و از اطلاعات این دادگان در آموزش طبقه بند MLKNN استفاده کنیم. ترکیب پیشنهادی نتایجی با دقت بسیار خوب بوده و تفاوت آن با روش‌های موجود معنی‌دار است. آزمون این ترکیب بر روی مجموعه داده تخصصی Yeast, core15k, و Scene نشان می‌دهد توانایی روش پیشنهادی ما با تعداد اندک دادگان برچسب‌دار بسیار بهتر از روش‌های دیگر است.

ساختار ادامه مقاله بدین صورت است که در فصل دو مروری بر کارهای مرتبط قبلی خواهیم داشت. در فصل سه روش پیشنهادی را ارائه می‌کنیم و در فصل چهارم به شرح دادگان، آزمایش‌های انجام‌گرفته و نتایج می‌پردازیم. نهایتاً در فصل پنجم به بررسی و جمع‌بندی می‌پردازیم.

۳ کارهای پیشین

در این بخش مروری بر کارهای انجام‌شده در حوزه یادگیری نیمه نظارتی مبتنی بر گراف خواهیم داشت. در روش یادگیری نیمه نظارتی مبتنی بر گراف ما دادگان را به صورت گرافی مدل می‌کنیم که منیفلدی^۸ از دادگان است. اکثر روش‌های مبتنی بر گراف به دنبال تخمین تابع پیش‌بینی برچسب بر روی گراف هستند که در تخمین این تابع باید دو فرض اساسی روش‌های یادگیری نیمه نظارتی در نظر گرفته شود: ۱- تخمین برچسب‌ها توسط تابع برای گره‌های برچسب‌دار به برچسب آن‌ها نزدیک باشد. ۲- تابع بر روی گراف هموار^۹ باشد. معمولاً این دو فرض را به عنوان یک چارچوب^{۱۰} در نظر می‌گیرند که بخش اول آن به عنوان تابع خطا و بخش دوم به عنوان تنظیم‌کننده^{۱۱} بیان می‌شود.

در روش ارائه‌شده در [۷] با استفاده از تعریف خطای مربعی و ماتریس لاپلاسی^{۱۲} به عنوان عبارت تنظیم‌کننده نمونه‌های بدون برچسب را برچسب‌زنی کرده‌اند. در [۸] با استفاده از لاپلاسی نرمال شده بجای لاپلاسی معمولی و ارائه روشی نیمه نظارتی الگوریتمی جدید جهت انتشار برچسب ارائه شده است. این دو روش اساساً اغلب روش‌های دیگر در این حوزه هستند. در [۵].

[۱۰] نمونه‌های بدون برچسب در کلاس جدیدی قرار می‌گیرند.

کلماتی که محتوای تصویر را توصیف کنند، اطلاق می‌شود. برچسب‌زنی تصاویر می‌تواند به صورت دستی یا خودکار انجام گیرد. در روش دستی، برچسب‌زنی تصاویر توسط افراد خیره انجام می‌پذیرد. به دلیل رشد سریع تعداد تصاویر تولیدشده در صفحات وب و شبکه‌های اجتماعی، استفاده از این روش بسیار وقت‌گیر و پرهزینه است. از این رو روش‌های برچسب‌زنی خودکار تصاویر معرفی شده‌اند. در برچسب‌زنی خودکار تصاویر یک تصویر می‌تواند با چندین برچسب مانند «جاده»، «اتومبیل» و «درخت» برچسب زده شود که از سه طبقه مختلف می‌باشند. در حوزه یادگیری ماشین به این نوع مسائل، مسائل یادگیری چندبرچسبی^۴ گفته می‌شود که کاربردهای مختلفی در دسته‌بندی صفحات وب، دسته‌بندی متون و برچسب‌زنی خودکار تصاویر دارد [۱].

الگوریتم‌های یادگیری بانظر در مسائل طبقه‌بندی موفق عمل می‌کنند اما موفقیت آنها وابسته به نمونه‌های برچسب‌دار است تا با آموزش بهتر بتوانند با دقت بالایی عمل نمایند. معمولاً جمع‌آوری این نمونه‌ها پرهزینه و وقت‌گیر است. برای حل این مشکل روش‌های یادگیری نیمه نظارتی در سال‌های اخیر مورد توجه پژوهشگران قرار گرفته است. در این روش، یادگیری از روی تعداد اندکی نمونه دارای برچسب و تعداد زیادی نمونه بدون برچسب انجام می‌گیرد. روش‌های یادگیری نیمه نظارتی مختلفی ارائه شده است که در [۲] مرور مناسبی به آن‌ها انجام یافته است. در سال‌های اخیر بیشترین فعالیت‌ها و پژوهش‌های این حوزه بر روی روش‌های یادگیری نیمه نظارتی مبتنی بر گراف انجام پذیرفته است [۳-۵]. این روش‌ها معمولاً دادگان را به صورت یک گراف مدل می‌کنند که در آن هر رأس گراف یک نمونه برچسب‌دار یا بدون برچسب است و یال بین دو رأس نشان‌دهنده میزان شباهت بین آنهاست. هدف آنها برچسب‌زنی رئوس بدون برچسب از روی رئوس دارای برچسب است [۶-۸]. روش یادگیری چندبرچسبی^۵ ML-KNN [۹] که در حوزه یادگیری بانظر ارائه شده است، به نوعی مبتنی بر گراف عمل می‌کند و بر اساس محاسبه احتمال MAP در بین همسایگان، نمونه جدید را برچسب‌زنی می‌کند.

با توجه به اینکه اغلب کارهای انجام‌گرفته در حوزه یادگیری نیمه نظارتی مبتنی بر گراف بر روی مسائل تک برچسبی متمرکز شده‌اند [۵، ۱۰-۱۱]. ما در این مقاله با ترکیب روش یادگیری نیمه نظارتی LGC^۶ [۶] و روش ML-KNN [۹] یک روش جدید بنام LGC+ML-KNN^۷ ارائه داده‌ایم. ما با استفاده از روش نیمه نظارتی نمونه‌های بدون برچسب را برچسب‌زنی اولیه می‌کنیم و در آموزش طبقه بند ML-KNN مشارکت می‌دهیم. بر این اساس، می‌توانیم با استفاده از مقادیر برچسب‌گره‌های همسایه مقادیر بهتری را برای نمونه‌های دیگر به دست آورده و نمونه‌های

8 Manifold

9 Smooth

10 Framework

11 Regulator

12 Laplacian

4 Multi-label learning

5 Multi-Label K-Nearest Neighbors

6 Local and Global Consistency

7 Local and Global Consistency+ Multi-Label K-Nearest Neighbors

می‌کند. روشی مشابه در [۱۶] با استفاده ماقربین لاپلاسیان و گسترش جمله تنظیم‌کننده مسئله چندبرچسبی را مدل می‌کند. جدول ۱ مشخصات اصلی کارهای پیشین را به صورت خلاصه نمایش می‌دهد. به‌طور کلی می‌توان گفت که کارهای پیشین در اهداف و نوع عملکرد باهم متفاوت هستند. در حقیقت هر یک از آنها زیرمجموعه‌ای از مفاهیم «یادگیری نیمه نظارتی»، «معیارهای سنجش شباهت»، «مدل‌سازی مبتنی گراف» «بازیابی تک برچسبی»، «بازیابی چند کلاسی»، «بازیابی چندبرچسبی» و «تبدیل صورت مساله به طبقه‌بندی» را مورد توجه قرار داده‌اند. با توجه به تنوع این اهداف مقایسه کارها با هم واقعا مشکل است. با وجود اینکه که بسیاری از مسائل در دنیای واقعی به صورت چندبرچسبی هستند؛ اغلب روش‌های یادگیری نیمه نظارتی مبتنی بر گراف بر روی مسائل تک برچسبی مطرح و پیاده‌سازی شده‌اند. همچنین از آنجایی که حجم داده دارای برچسب با دقت سیستم یادگیری شده رابطه مستقیم دارد، به نظر می‌رسد ترکیب روش‌های یادگیری نیمه نظارتی و ترالقای^{۱۷} می‌تواند با افزایش حجم دادگان دارای برچسب دقت برچسب‌زنی را بهبود بخشد.

۴ روش ارائه شده

در این بخش روش پیشنهادی را معرفی می‌کنیم. روش پیشنهادی ما در عین اینکه ماهیت نیمه نظارتی و چندبرچسبی دارد؛ از ترکیب روش‌های دیگری که تک برچسبی هستند به‌عنوان گام اول استفاده می‌کند. ما با ترکیب الگوریتم یادگیری نیمه نظارتی مبتنی بر گراف (LGC) و الگوریتم یادگیری چندبرچسبی -ML-KNN روشی ترکیبی و کارا برای این‌گونه مسائل ارائه داده‌ایم که در مقایسه با روش‌های موجود نتایج بهتری را به دست می‌دهد. پیش از ورود به بحث روش پیشنهادی، برخی مقدمات لازم برای درک مفاهیم مورد استفاده را توضیح می‌دهیم. همچنین با توجه به اینکه روش پیشنهادی بر اساس دو روش LGC و ML-KNN ارائه شده است هر دو روش را به تفصیل بررسی می‌کنیم.

۴-۱ یادگیری نیمه نظارتی

در یادگیری ماشین به مسائلی که فرآیند یادگیری در آن‌ها با استفاده از نمونه‌های برچسب‌دار انجام می‌گیرد یادگیری بانظر یا بانظارت گفته می‌شود که در آن هدف سیستم یادگیر به دست آوردن فرضیه‌ای است که تابع یا رابطه بین ورودی و خروجی را حدس بزند. در یادگیری بدون نظارت یادگیری از روی داده‌های بدون برچسب انجام می‌گیرد. هدف از یادگیری، یافتن الگوهای پنهان در این داده‌ها است. یادگیری نیمه نظارتی^{۱۸} دسته‌ای از روش‌های یادگیری ماشین است که در آن از داده‌های بدون برچسب و

رابطه همسایگی بین نمونه‌ها به کمک روش تعبیه خطی محلی^{۱۳} تعریف می‌شود و بر اساس آن برچسب‌زنی صورت می‌گیرد. این روش‌ها برای نمونه‌های تک برچسبی کارایی دارند. در [۴] با ساخت یک مدل مولد^{۱۴} برای نمونه‌های تک برچسب و استفاده از الگوریتم EM^{۱۵} عمل برچسب‌زنی انجام می‌گیرد.

جدول ۱ خلاصه مشخصات کارهای پیشین

مرجع	دادگان	نوع	معیار ارزیابی	تابع شباهت
Wang [۳]	Yeast Core5k Scene	چند برچسبی	Average precision Micro-F1	هسته گوسی
Mehrkanoon [۲۱]	Yeast Image Scene Emotion	چند برچسبی	Average precision Micro-F1 Ranking Loss Hamming Loss	هسته گوسی
Zhao [۵]	Yale-B MNIST USPS ETH80 CPOIL100	تک برچسبی	Average Accuracy	LLE
Chen [۱۵]	subset of RCV1-v2	چند برچسبی	Micro-F1	هسته گوسی
Zha [۱۶]	TRECVID 2006		MAP	هسته گوسی
Zhou [۶]	USPS20-newsgroups dataset	تک برچسبی	Test error	هسته گوسی

در زمینه یادگیری نیمه نظارتی چندبرچسبی تلاش‌های چندانی صورت نگرفته اما یکی از روش‌های معمول، روش‌های مبتنی بر رتبه‌بندی برچسب است [۱۲] که با یادگیری یک تابع رتبه‌بند از روی نمونه‌های برچسب‌دار و با اعمال یک حد آستانه و امتیازدهی، نمونه جدید را برچسب می‌زند. این روش با وجود مقیاس‌پذیری خوب، در بهره‌گیری از وابستگی بین برچسب‌ها ضعیف عمل می‌کند. در [۱۳] به وابستگی بین برچسب‌ها توجه شده و با ارائه یک روش مبتنی بر آنتروپی حداکثر، مدلی سلسله مراتبی ارائه گردیده است. در [۱۴] روشی برای یادگیری چندبرچسبی ارائه شده که صریحاً از اطلاعات کلاس‌های مختلف و وابستگی آن‌ها استفاده می‌کند. یکی دیگر از روش‌های یادگیری چندبرچسبی نیمه نظارتی [۱۵] با گسترش جمله تنظیم‌کننده به دو بخش و با استفاده از حل معادله سیلوستر^{۱۶} نمونه‌های بدون برچسب را برچسب‌زنی

13 Locally Linear Embedding (LLE)

14 Generative Model

15 Expectation Maximization

16 Sylvester Equation

۱۷ معادل فارسی دیگر پیشنهادی برای transductive واژه ورارسایی است. واژه ترالقای هم خاصیت القا و هم تراگذاری را به ذهن می‌آورد.
18 Semi-supervised Learning

Archive of SID

Algorithm $Y_u = LGC(V, E, Y_l)$

Input: Graph (V, E) , labels Y_l

Output: labels Y_u

$$\forall_i, D_{ii} = \sum_j W_{i,j}$$

$$S = D^{-1/2} W D^{-1/2}$$

$$Y_u = 0$$

$$Y^{(0)} = [Y_l, Y_u], t = 0$$

Repeat

$$Y^{(t+1)} = \alpha S Y^t + (1 - \alpha) Y^0$$

$$t = t + 1$$

Until $Y^{(t)}$ Converges

الگوریتم ۱ الگوریتم یادگیری چندبرجسی LGC [۶]

Algorithm $[y_t, r_t] = ML-KNN(T, K, t, s)$

%Computing the prior probabilities $P(H_b^l)$

(1) for $l \in y$ do

$$(2) p(H_1^l) = (s + \sum_{xi=1}^m y_{xi}(l)) / (2s + m); p(H_0^l) = 1 - p(H_1^l)$$

%Find Posterior probabilities $P(E_j^l | H_b^l)$

(3) Identify $N(x_i), i \in \{1, 2, \dots, m\};$

(4) for $l \in y$ do

(5) for $j \in \{0, 1, \dots, k\}$ do

$$(6) c[j] = 0; c'[j] = 0;$$

(7) for $i \in \{1, 2, \dots, m\}$ do

$$(8) \delta = C_{xi}(l) = \sum_{a \in N(x_i)} y_a(l)$$

(9) if $(y_{xi}(l) = 1)$ then $c[\delta] = c[\delta] + 1;$

(10) else $c'[\delta] = c'[\delta] + 1;$

(11) for $j \in \{0, 1, \dots, k\}$ do

$$(12) P(E_j^l | H_1^l) = (s + c[j]) / (s(k+1) + \sum_{p=0}^k c[p])$$

$$(13) P(E_j^l | H_0^l) = (s + c'[j]) / (s(k+1) + \sum_{p=0}^k c'[p])$$

%Computing y_t, r_t

(14) Identify $N(t);$

(15) for $l \in y$ do

$$(16) C_t(l) = \sum_{a \in N(t)} y_a(l);$$

$$(17) y_t(l) = \arg \max_{b \in \{0,1\}} P(H_b^l) P(E_{C_t(l)}^l | H_b^l);$$

$$(18) r_t(l) = \frac{P(H_1^l) P(E_{C_t(l)}^l | H_1^l)}{\sum_{b \in \{0,1\}} P(H_b^l) P(E_{C_t(l)}^l | H_b^l)}$$

الگوریتم ۲- الگوریتم چندبرجسی ML-KNN [۹]

۳-۴ الگوریتم چندبرجسی ML-KNN

(ML-KNN) الگوریتم طبقه‌بندی چندبرجسی بر مبنای K

همسایه نزدیک هست [۹]. این الگوریتم نظارتی، با به دست آوردن

داده‌های برجسب‌دار به صورت هم‌زمان برای یادگیری استفاده می‌شود و هدف این روش مشارکت دادن دادگان بدون برجسب به همراه دادگان برجسب‌دار اندک در فرآیند آموزش است که باعث افزایش دقت تصمیم‌گیری می‌شود.

۲-۴ یادگیری نیمه نظارتی مبتنی بر گراف

الگوریتم LGC یکی از روش‌های یادگیری نیمه نظارتی مبتنی بر گراف است که از ماتریس لاپلاسیین نرمال گراف برای برجسب‌زنی نمونه‌های بدون برجسب استفاده می‌کند. در این روش ابتدا ماتریس شباهت W و سپس از روی آن ماتریس قطری D و لاپلاسیین نرمال گراف محاسبه می‌شوند. مقادیر این ماتریس اطلاعات مفیدی در رابطه با طیف و ساختار گراف در اختیار ما قرار می‌دهد. بر اساس اصل پیوستگی و فرض‌های یادگیری نیمه نظارتی مبتنی بر گراف می‌توان گفت که اگر نقاطی از نظر طیفی شبیه به هم باشند و در یک ناحیه چگال قرار داشته باشند، برجسب آن‌ها نیز مشابه یکدیگر است [۶].

در گراف وزن‌دار $G = (V, E)$ ، V مجموعه رئوس متناظر با هرکدام از نمونه‌ها هست. E مجموعه یال‌های بین رئوس بوده و با ماتریس شباهت W نشان داده می‌شود. هر درآیه ماتریس W با تابع هسته گوسی از رابطه (۱) محاسبه می‌شود:

$$w_{ij} = \exp(-||x_i - x_j|| / \sqrt{2s^2}) \quad (1)$$

در این رابطه x_i و x_j ، نمونه‌های دارای برجسب و بدون برجسب می‌باشند و $d(x_i, x_j)$ ، فاصله اقلیدسی این دو است. s انحراف معیار فواصل است که از روی مجموع دادگان محاسبه می‌شود. مجموعه دادگان آموزشی $X = \{X_l, X_u\}$ از ترکیب دو مجموعه جداگانه داده‌های برجسب‌دار $X_l = \{X_{l1}, X_{l2}, \dots, X_{lk}\}$ و بدون برجسب $X_u = \{X_{u1}, X_{u2}, \dots, X_{un}\}$ تشکیل شده است؛ و مجموعه برجسب $Y_{nc} = \{Y_l, Y_u\}$ که n تعداد نمونه‌ها و c تعداد کلاس برجسب‌ها هست. Y_l ماتریس برجسب برای X_l و Y_u ماتریس برجسب برای X_u هست. $Y_l(i, k) = 1$ اگر X_i دارای برجسب کلاس k ام باشد و در غیر این صورت برابر صفر است. ماتریس Y_u در ابتدا برابر صفر هست که هدف پیش‌بینی برجسب برای مجموعه Y_u هست. با استفاده از الگوریتم LGC پیش‌بینی برجسب برای Y_u را انجام می‌دهیم. در این الگوریتم ماتریس پیش‌بینی برجسب Y که متشکل از Y_l ماتریس برجسب برای X_l و Y_u که در ابتدا با مقدار صفر مقداردهی می‌شود و در طی یک فرآیند تکراری و تا زمان همگرایی مقادیر کلاس برجسب‌ها، مقدار ماتریس Y بدین صورت محاسبه می‌شود که مقادیر Y_l بدون تغییر در گام بعدی مشارکت داده می‌شوند و مقادیر Y_u از حاصل ضرب مقادیر ماتریس لاپلاسیین نرمال گراف در ماتریس برجسب Y به دست می‌آیند. پارامتر α میزان تاثیرگذاری Y_l در محاسبه Y_u را کنترل می‌کند. (الگوریتم ۱)

بندی برچسب ها از یک حد آستانه، مثلا ۰.۵، استفاده می شود و برچسب های بالاتر از آن مقدار یک و مقادیر کمتر صفر منظور می گردد. ما در روش پیشنهادی خود با توجه به کاردینالیتی هر مجموعه داده مقادیر با بیشترین احتمال را از r_i استخراج کرده و بردار برچسب y_i را به دست می آوریم.

Algorithm Y=LGC+ML-KNN(X,Y_l, k, t)

%Transductive step:

X = [X_l, X_u]; %X_l: labeled, X_u: unlabeled samples

Y = LGC(X, Y_l)

%Improvement step:

Y_u = ML-KNN(X, k, X_u, s)

الگوریتم ۳ - ترکیب الگوریتم های LGC+ML-KNN

۵ آزمایش ها و نتایج

در این بخش روش پیشنهادی خود را بر روی مجموعه داده های موجود و به دو روش اجرا می کنیم. پیش از بررسی نتایج با توجه به استفاده از تحلیل معنی داری آماری، مفاهیم مربوط به فاصله اطمینان و معناداری را در ابتدا معرفی می کنیم.

۵-۱ فاصله اطمینان و معنی داری

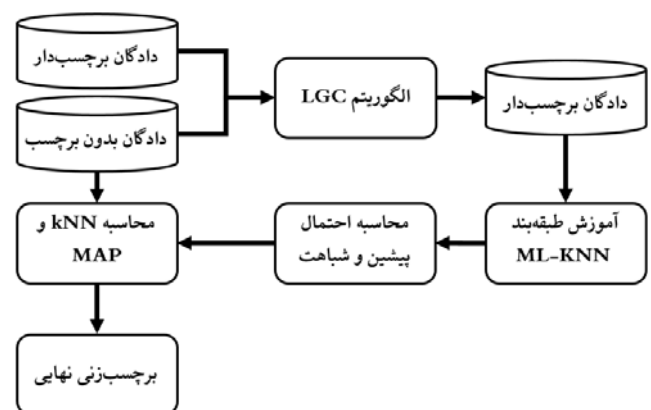
اگر X متغیری تصادفی دارای توزیع نرمال با میانگین m_x و انحراف معیار s_x باشد، می توان نواحی مختلفی از دامنه x را شناسایی کرد که احتمال تعلق x به آنها زیاد است. بازه های اطمینان از جمله این نواحی هستند؛ مثلا اگر x دارای توزیع نرمال با میانگین صفر و واریانس یک باشد و از این توزیع ۱۰۰ نمونه با شرایط یکسان به صورت تصادفی برداریم، انتظار داریم بیش از ۹۵ نمونه در بازه $(-2, 2)$ قرار داشته باشند. به صورت کلی بازه اطمینان a درصدی x که به صورت (X_u, X_l) نمایش داده می شود؛ به این معنی است اگر n نمونه به صورت تصادفی و با شرایط یکسان از توزیع x برداریم، $a \times n$ نمونه در این بازه جای می گیرند. X_l را کران پایین و X_u را کران بالای بازه می نامند. برای توزیع نرمال معمول این است که بازه ها به صورت متقارن (نسبت به میانگین) و با ضرایب صحیح انحراف معیار تعریف می شوند. بازه اطمینان $m_x \pm s_x$ ، $m_x \pm 2s_x$ ، $m_x \pm 3s_x$ به ترتیب با پوشش ۶۵ درصد، ۹۵ درصد و ۹۹ درصد سطح زیر منحنی توزیع نرمال شهرت بیشتری دارند.

در یادگیری ماشین و شناسایی الگو معمولا بازه اطمینان ۹۵ درصدی برای مطالعه محدوده متغیرها استفاده می شود؛ مثلا دقت یک طبقه بند با عبارت 74 ± 3 بیان می شود. در این صورت بازه اطمینان دقت طبقه بند (۷۱ و ۷۷) درصد است یعنی اگر تعداد بسیاری زیادی آزمایش با این طبقه بند انجام شود، انتظار می رود که در اغلب موارد (یعنی بیش از ۹۵ درصد موارد) دقت آزمایش در محدوده ۷۱ تا ۷۷ درصد باشد. اگر میانگین دقت طبقه بند خاص دیگری ۷۸ درصد باشد، گفته می شود که طبقه بند دوم

مقدار MAP^{۱۹} از روی احتمال پیشین و فراوانی هر کلاس در K همسایه نزدیک، مجموعه آزمون را برچسب زنی می کند. (الگوریتم ۲). در این الگوریتم، ورودی مجموعه آموزشی T ، K نشان دهنده تعداد همسایه ها و t نمونه آزمون و s پارامتر هموارسازی است. خروجی الگوریتم y_i که بردار دسته بندی نمونه آزمون را مشخص می کند و r_i که بردار برچسب با مقادیر واقعی است که رتبه برچسب از روی این مقادیر و متناسب با تعداد برچسب های مورد نیاز حاصل می شود. در گام (۱) و (۲) مقدار احتمال پیشین $P(H_b^i)$ برای تمام کلاس برچسب ها محاسبه می شود. گام (۳) تا (۱۳) مقدار احتمال پسین $P(E_j^i | H_b^i)$ را محاسبه می کند که $c[j]$ در بین k همسایه نزدیک به ازای هر کلاس برچسب l تعداد نمونه های آموزشی دارای برچسب l را محاسبه می کند و در نهایت در گام (۱۴) تا (۱۸) بردار دسته بندی کلاس برچسب خروجی برای نمونه آزمون t محاسبه می شود. ما برای پیاده سازی این بخش از پیاده سازی نویسندگان اصلی [۹] استفاده کرده ایم که به صورت متن باز در دسترس است.^{۲۰}

۴-۴ ترکیب LGC و MLKNN

نمودار بلوکی روش ارائه شده در شکل ۱ نشان داده شده است. در این روش مطابق الگوریتم ۳ ابتدا با استفاده از روش LGC، نمونه های بدون برچسب را برچسب زنی می کنیم تا بتوانیم از این نمونه ها نیز در فرآیند آموزش استفاده نماییم. سپس با استفاده از خروجی مرحله اول، طبقه بند ML-KNN را آموزش می دهیم که با توجه به مشارکت تعداد نمونه های بیشتر در فرآیند آموزش، طبقه بند بهتر آموزش می بیند و با به دست آوردن پارامترهای بهتر می توانیم نمونه های آزمون جدید یا نمونه های بدون برچسب قبلی را با دقت بالاتر و کیفیت بهتر برچسب زنی کنیم.



شکل ۱ نمودار بلوکی روش پیشنهادی

در بخش پایانی الگوریتم از روش کاردینالیتی مجموعه داده ها استفاده و بهبود نتایج را مشاهده کردیم. در ML-KNN برای رتبه

19 Maximum a Posterior

20 <http://cse.seu.edu.cn/PersonalPage/zhangml/>

Archive of SID

۱,۵۷±۴,۲۴ برچسب خورده است و برای مطالعه آن‌ها و پروتئین‌های آنها بکار می‌رود [۱۹].

Scene: این پایگاه داده تصویری شامل ۲۴۰۷ نمونه با ۲۹۴ ویژگی هست که ۶ کلاس برچسب دارند و هر نمونه به‌طور متوسط ۱,۰۷ برچسب کلاس دارد.

۳-۵ معیارهای ارزیابی

نحوه ارزیابی کارایی سیستم‌های یادگیری چندبرچسبی با سیستم‌های تک‌برچسبی متفاوت است. معیارهای ارزیابی عمومی در سیستم‌های تک‌برچسبی شامل صحت^{۲۳}، دقت^{۲۴}، بازخوانی^{۲۵} و معیار^{۲۶} است [۲۰]. در سیستم‌های یادگیری چندبرچسبی ارزیابی کمی پیچیده‌تر هست، برای ارزیابی نتایج از معیارهای ارزیابی زیر استفاده کرده‌ایم (جدول ۳) [۲۱]:

میانگین دقت^{۲۷}: جزو معیارهای مبتنی بر رتبه^{۲۸} و در اصل یکی از معیارهای ارزیابی بازایی اسناد مرتبط در سیستم‌های بازایی اطلاعات هست. میانگین کسری از رتبه برچسب‌های بالاتر از یک برچسب خاص $y \in Y$ را که در Y هستند را ارزیابی می‌کند. این معیار هر چه قدر به یک نزدیک‌تر باشد، بهتر است.

Hamming Loss: جزو معیارهای مبتنی بر نمونه^{۲۹} هست و تفاوت مقارن دو مجموعه برچسب را محاسبه می‌کند. بررسی می‌کند که چند بار یک زوج برچسب نمونه، به اشتباه جزء برچسب طبقه‌بندی شده است. بدین معنی که برچسبی که متعلق به نمونه نیست به‌عنوان برچسب پیش‌بینی شده و برچسبی که متعلق به نمونه بوده پیش‌بینی نشده است. مقادیر پایین این معیار مطلوب است یعنی این معیار هر چه قدر نزدیک به صفر باشد بهتر است.

Ranking Loss: جزء معیارهای مبتنی بر رتبه هست و متوسط کسری از زوج برچسب‌ها را که به‌طور معکوس برای نمونه مرتب‌شده‌اند را ارزیابی می‌کند و هر چه به صفر نزدیک‌تر باشد بهتر خواهد بود.

MicroF1: جزو معیارهای مبتنی بر برچسب^{۳۰} هست و همانند معیار^{۳۱}، اما از روی Micro Precision و Micro Recall به دست می‌آید.

به‌صورت معنی‌داری دقیق‌تر از طبقه بند اول عمل کرده است. همچنین اگر میانگین دقت آن کمتر از ۷۱ درصد باشد؛ گفته می‌شود که طبقه بند دوم به‌صورت معنی‌داری ضعیف‌تر از طبقه بند اول عمل کرده است. نهایتاً اگر میانگین دقت طبقه بند دوم بین ۷۱ تا ۷۷ (مثلاً ۷۶ درصد) باشد گفته می‌شود تفاوت دقت دو طبقه‌بند ظاهری بوده و معنی‌دار نیست؛ یعنی بر اساس این تفاوت نمی‌توان به برتری طبقه بند دوم رای داد. اگر تعداد نمونه‌های بکار رفته در آزمایش ۳۰ نمونه یا بیشتر باشد؛ می‌توان بر اساس قضیه حد مرکزی^{۳۱} فرض کرد که نتایج آزمایش توزیع نرمال دارد [۱۷]. بر این اساس فاصله اطمینان ۹۵ درصدی p از رابطه زیر محاسبه می‌شود:

$$C = p \pm 1.96 \sqrt{\frac{p(1-p)}{N}} \quad (2)$$

در این رابطه n تعداد نمونه‌های آزمون، p میانگین کمیت و C فاصله اطمینان آن است. اگر میانگین نتایج یک آزمایش خارج از فاصله اطمینان نتایج آزمایش دیگر باشد، آنگاه تفاوت نتایج این دو آزمایش معنی‌دار تلقی می‌گردد. در این پژوهش نیز از این روش برای بررسی و تحلیل نتایج به دست آمده استفاده می‌کنیم.

۲-۵ دادگان

ما در این بخش روش پیشنهادی خود را بر روی سه مجموعه داده موجود که برای مسائل یادگیری چندبرچسبی تصاویر بکار می‌روند اجرا می‌کنیم. مشخصات این سه در جدول ۲ نشان داده شده است.

جدول ۲ مشخصات مجموعه دادگان مورد استفاده

dataset	#instance	#attribute	#label	l.c	Domain
Corel5k	4834	500	43	3.5	image
Yeast	2417	103	14	4.2	biology
Scene	2407	294	6	1.07	image

l.c: label cardinality

Corel 5k: مجموعه داده اصلی Corel 5k که در [۱۸] استفاده شده، دارای ۵۰۰۰ نمونه هست که هر تصویر به سه تا ده ناحیه تقسیم شده است این نواحی به ۵۰۰ خوشه خوشه‌بندی شده‌اند که هر کدام blob^{۳۲} نامیده می‌شوند. هر تصویر به‌صورت یک بردار باینری از این ۵۰۰ خوشه نشان داده می‌شوند. هر تصویر این مجموعه داده به‌طور میانگین با ۳,۵ واژه برچسب خورده است. به تاسی از کارهای مشابه، واژگانی را که کمتر از ۱۰۰ بار استفاده شده‌اند کنار گذاشتیم. در نهایت ۴۳ واژه و ۴۸۳۴ تصویر را استفاده کردیم.

Yeast: این مجموعه شامل ۲۴۱۷ نمونه با ۱۰۳ ویژگی هست که ۱۴ کلاس برچسب دارند و هر نمونه به‌طور متوسط با

- 23 Accuracy
- 24 Precision
- 25 Recall
- 26 F-measure
- 27 Average Precision
- 28 Ranking Based
- 29 Example Based
- 30 Label Based

- 21 Central limit theorem
- 22 Binary Large Object (BLOB)

جدول ۳ معیارهای ارزیابی نتایج

Average Precision; $AP = \frac{1}{N} \left(\sum_{i=1}^N (1/i) \sum_{j=1}^i y_j \right)$
Hamming loss; $HL = \frac{1}{N} \sum_{i=1}^N (1/Q) (h(x_i) \Delta y_i)$
Ranking loss; $RL = \frac{1}{N} \sum_{i=1}^N \frac{ D_i }{ y_i \cdot y_l }$
Micro Average; $Micro_{AP} = \frac{\sum_{j=1}^N tp_j}{\sum_{j=1}^N tp_j + \sum_{j=1}^N fp_j}$
Micro Recall, $Micro_{Re} = \frac{\sum_{j=1}^N tp_j}{\sum_{j=1}^N tp_j + \sum_{j=1}^N fn_j}$
$Micro_{F1} = \frac{2 \times Micro_{Pr} \times Micro_{Re}}{Micro_{Pr} + Micro_{Re}}$

N: تعداد نمونه‌ها، y_i : برچسب نمونه i ام، $| \cdot |$: عملگر اندازه (طول، تعداد) Δ : عملگر تفاضل متقارن دو مجموعه. tp =true positive, fp =false negative
 positive, fn =false negative
 برچسب پیش‌بینی شده برای نمونه x توسط طبقه بند h

۴-۵ نتایج

نتایج آزمون روش خود را در مقایسه با روش (dlp) ارائه شده در [۳] و روش (FS-MLSS-KSC) ارائه شده در [۲۲] و سایر روش‌های موجود قبلی که در این دو مقاله مقایسه شده‌اند بررسی کرده‌ایم. روش آزمون در [۳] بدین شکل هست که بر روی مجموعه داده‌های Yeast، corel 5k و Scene، با انتخاب نرخ برچسب‌های مختلف (به‌طور مثال با نرخ ۰,۰۲) دو درصد از نمونه‌ها را به‌عنوان دادگان برچسب‌دار انتخاب می‌کنیم و نیمی از مابقی نمونه‌ها را به‌عنوان دادگان بدون برچسب در نظر می‌گیریم و با ده مرتبه اجرا برای هر نرخ، نتایج را مورد بررسی قرار می‌دهد و در [۲۲] با نسبت ۳۰ درصد دادگان برای آزمون و ۷۰ درصد برای آموزش که دادگان آموزش به دو بخش با برچسب و بدون برچسب تقسیم شده‌اند، نتایج حاصل را مورد بررسی قرار می‌دهد.

در آزمون‌های بازیابی اطلاعات معمول این است که مقادیر محض معیارهایی مانند میانگین دقت، بازنمایی و معیارهای رتبه‌ای را باهم مقایسه می‌کنند. دلیل این امر این است که تفاوت‌های بسیار جزئی این مقادیر معادل تفاوت‌های چشمگیر در ترکیب و تعداد نتایج بازیابی شده مرتبط است. با این وجود، تحلیل آماری نتایج می‌تواند نقاط قوت و ضعف سیستم‌ها را بهتر نشان دهد [۲۳]. در اینجا برای بررسی جزئیات نتایج از تحلیل معنی‌داری آماری^{۳۱} مبتنی بر فاصله یا بازه اطمینان^{۳۲} استفاده می‌کنیم.

جدول ۴ فاصله اطمینان میانگین دقت روش‌های مختلف را گزارش کرده است. نتایج آزمایش‌های ما نشان می‌دهد که روش ارائه شده (LGC+MLKNN) نسبت به روش‌های قبلی بهتر عمل می‌کند، بخصوص هنگامی که تعداد نمونه‌های برچسب خورده بسیار کم هست. به‌طور مثال در مجموعه داده Corel5k میانگین دقت به دست آمده توسط روش پیشنهادی 0.365 ± 0.014 است که نسبت به نتیجه روش dlp که دارای میانگین دقت با فاصله اطمینان 0.33 ± 0.013 است تفاوت معناداری را نشان می‌دهد.

جدول ۵ فاصله اطمینان روش پیشنهادی و سایر روش‌ها بر روی معیار MicroF1 را گزارش می‌کند که نشان می‌دهد نتایج روش ما بر روی مجموعه داده Corel5k با نرخ انتخاب برچسب ۰,۰۲ دارای نتیجه 0.261 ± 0.013 است که در مقایسه با بهترین نتیجه کارهای پیشین که مربوط به روش dlp با مقدار 0.24 ± 0.012 به‌طور معناداری بهتر عمل نموده و دقت روش ما در معیار MicroF1 از مقدار حداکثر فاصله اطمینان dlp بیشتر است. با انتخاب نرخ برچسب‌های بالاتر شاهد این بهبود نتایج به‌طور معناداری هستیم. نتایج روش پیشنهادی در دو مجموعه داده Scene و Yeast نیز دارای بهبود است. این نتایج همان‌طور که در جدول مشخص است نسبت به سه روش دیگر به‌طور معناداری بهبود دارد و نسبت به روش dlp نیز در نتایج بهتر و در بحث معناداری هم‌تراز عمل نموده‌ایم و نتایج ما در فاصله اطمینان روش dlp است.

جدول ۶ فاصله اطمینان روش‌های مختلف را برای برچسب‌زنی در حالت استقرایی نشان می‌دهد. هدف از پیاده‌سازی روش پیشنهادی در این حالت این است که عملکرد آن را در حالتی که بخشی از دادگان به‌عنوان نمونه آزمون در نظر گرفته می‌شوند و در فرآیند آموزش دخالت ندارند را بررسی کنیم. این جدول نشان می‌دهد که نتایج روش‌های مختلف بر روی سه مجموعه داده کاملاً باهم متفاوت است. اگر این سه مجموعه داده را به دودسته دادگان تصویری مشتمل بر Corel5k و Scene و دادگان غیر تصویری Yeast تقسیم کنیم می‌توانیم نتیجه‌گیری کنیم که روش پیشنهادی بر روی دادگان تصویری عملکردی مناسب و قابل قبول داشته است. بخصوص با نگاهی به فاصله‌های اطمینان گزارش شده در این جدول می‌توان دید که بر روی مجموعه دادگان Scene نتایج روش پیشنهادی به نحوی معنی‌داری بهتر از نتایج سایر روش‌هاست.

31 Statistical Significance Analysis

32 Confidence Interval

جدول ۶ مقایسه برچسب‌زنی روش پیشنهادی و روش‌های موجود

	RL	AP	HL	Micro_F1
Corel5k, labeled: 500, unlabeled: 900, test 600				
FS-MLSS-KSC	0.173±0.03	0.789±0.033	0.168±0.03	0.591±0.039
TRAM	0.2±0.032	0.751±0.035	0.201±0.032	0.553±0.04
ML-LOC	0.171±0.03	0.682±0.037	0.176±0.03	0.548±0.04
ML-KNN	0.209±0.033	0.75±0.035	0.205±0.032	0.35±0.038
Proposed	0.184±0.031	0.789±0.033	0.171±0.03	0.618±0.039
Scene, labeled: 500, unlabeled: 711, test 1196				
FS-MLSS-KSC	0.477±0.028	0.576±0.028	0.199±0.023	0.42±0.028
TRAM	0.514±0.028	0.552±0.028	0.213±0.023	0.386±0.028
ML-LOC	0.497±0.028	0.516±0.028	0.195±0.022	0.403±0.028
ML-KNN	0.522±0.028	0.531±0.028	0.221±0.024	0.342±0.027
Proposed	0.105±0.017	0.832±0.021	0.106±0.017	0.696±0.026
Yeast, labeled: 500, unlabeled: 1000, test 917				
FS-MLSS-KSC	0.18±0.025	0.751±0.028	0.199±0.026	0.626±0.031
TRAM	0.181±0.025	0.744±0.028	0.218±0.027	0.633±0.031
ML-LOC	0.338±0.031	0.712±0.029	0.199±0.026	0.619±0.031
ML-KNN	0.182±0.025	0.74±0.028	0.207±0.026	0.595±0.032
Proposed	0.178±0.025	0.742±0.028	0.211±0.026	0.642±0.031

۶ بحث و بررسی

در این مقاله در حوزه برچسب‌زنی تصاویر، به کمک روش‌های یادگیری نیمه نظارتی و الگوریتم‌های یادگیری چندبرچسبی، یک روش ترکیبی ارائه دادیم که توانایی برچسب‌زنی تصاویر را با دقت و صحت بالاتری نسبت به روش‌های موجود را دارد، با توجه به نتایج آزمایش‌ها، روش ما کارایی بهتری بر روی مجموعه داده‌های تخصصی تصاویر مانند Corel5k, Scene, image دارد و بر روی مجموعه داده Yeast که جزء دیتاست‌های بیولوژی است نیز کارایی قابل قبول و هم‌تراز روش‌های دیگر دارد.

با توجه به اینکه یکی از چالش‌های مهم در یادگیری نیمه نظارتی نحوه استفاده از دادگان بدون برچسب در فرآیند آموزش سیستم است، ما در مقاله حاضر توانسته‌ایم با استفاده از مدل یادگیری نیمه نظارتی مبتنی بر گراف دادگان بدون برچسب را در فرآیند آموزش دخالت دهیم و با دقت خوبی برچسب اولیه‌ای برای دادگان بدون برچسب پیش‌بینی کنیم و از اطلاعات این دادگان در آموزش طبقه‌بند MLKNN استفاده کنیم که این ترکیب نتایج با دقت بسیار خوب و معنی‌داری بر روی مجموعه داده تخصصی corel5k و Scene نشان می‌دهد چراکه ساختار روش MLKNN به نحوی است که بر اساس محاسبه احتمال برچسب‌ها در همسایگان نزدیک هر نمونه کار می‌کند و روش ما توانسته دقت گام اول را بهبود ببخشد. توانایی سیستم پیشنهادی ما هنگامی که دادگان بدون برچسب کم است بسیار بهتر از روش‌های دیگر است؛ یعنی با همین تعداد دادگان روش MLKNN و یا

جدول ۴ مقایسه معیار AP روش پیشنهادی و روش‌های موجود

Label rate	0.02	0.04	0.06	0.08
Corel5k				
ML-KNN	0.29±0.013	0.298±0.013	0.315±0.014	0.316±0.014
TRAM	0.29±0.013	0.3±0.013	0.33±0.014	0.34±0.014
Boostext	0.26±0.012	0.27±0.013	0.292±0.013	0.31±0.014
dlp	0.33±0.013	0.34±0.014	0.36±0.014	0.378±0.014
Proposed	0.365±0.014	0.368±0.014	0.374±0.014	0.366±0.014
Scene				
ML-KNN	0.658±0.019	0.735±0.018	0.75±0.018	0.79±0.017
TRAM	0.734±0.018	0.773±0.017	0.81±0.016	0.82±0.016
Boostext	0.746±0.018	0.795±0.016	0.82±0.016	0.822±0.016
dlp	0.745±0.018	0.79±0.017	0.82±0.016	0.822±0.016
Proposed	0.76±0.017	0.793±0.017	0.816±0.016	0.823±0.016
ML-KNN	0.685±0.019	0.704±0.019	0.71±0.019	0.715±0.019
TRAM	0.705±0.018	0.721±0.018	0.725±0.018	0.73±0.018
Yeast				
Boostext	0.68±0.019	0.695±0.019	0.698±0.019	0.705±0.019
dlp	0.713±0.018	0.725±0.018	0.73±0.018	0.735±0.018
Proposed	0.714±0.018	0.711±0.018	0.727±0.018	0.733±0.018

جدول ۵ مقایسه (Micro F1) روش پیشنهادی و روش‌های موجود

Label rate	0.02	0.04	0.06	0.08
Corel5k				
ML-KNN	0.025±0.004	0.015±0.003	0.025±0.005	0.03±0.005
TRAM	0.18±0.011	0.2±0.012	0.22±0.012	0.235±0.012
Boostext	0.13±0.01	0.15±0.01	0.16±0.011	0.162±0.011
dlp	0.24±0.012	0.245±0.012	0.247±0.013	0.25±0.013
Proposed	0.261±0.013	0.258±0.013	0.277±0.013	0.276±0.013
Scene				
ML-KNN	0.25±0.017	0.4±0.02	0.512±0.021	0.57±0.021
TRAM	0.51±0.02	0.6±0.02	0.648±0.02	0.672±0.02
Boostext	0.508±0.02	0.595±0.02	0.645±0.02	0.675±0.02
dlp	0.54±0.02	0.63±0.02	0.65±0.02	0.675±0.02
Proposed	0.552±0.02	0.635±0.02	0.664±0.019	0.682±0.019
Yeast				
ML-KNN	0.535±0.02	0.55±0.02	0.56±0.02	0.578±0.021
TRAM	0.58±0.02	0.618±0.02	0.627±0.02	0.63±0.02
Boostext	0.55±0.02	0.56±0.02	0.558±0.02	0.568±0.021
dlp	0.598±0.02	0.62±0.02	0.63±0.02	0.638±0.02
Proposed	0.601±0.02	0.582±0.02	0.603±0.02	0.617±0.02

- X. Z. G. a. J. D. L. Zhu, "Semi-supervised learning using Gaussian fields and harmonic functions," *Proceedings of the Twentieth International Conference on Machine Learning (ICML)*, Washington, p. p.912-919, DC, 2003.
- [8] X. Z. G. a. J. D. L. Zhu, "Semi-supervised learning using gaussian fields and harmonic functions." , *Proceedings of the 20th International conference on Machine learning (ICML-03)*, 2003.
- [9] M.-L. a. Z.-H. Z. Zhang, "ML-KNN: A lazy learning approach to multi-label learning.," *Pattern recognition* 40.7, pp. 2038-2048., 2007.
- [10] F. e. a. Nie, "A general graph-based semi-supervised learning with novel class discovery." : ., *Neural Computing and Applications* 19.4, pp. 549-555, 2010.
- [11] L. a. G. Y. Feng, "Semi-Supervised Classification Based on Mixture Graph.," *Algorithms* 8.4, pp. 1021-1034., 2015.
- [12] K. a. Y. S. Crammer, "A new family of online algorithms for category ranking.," in *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval. ACM*, 2002.
- [13] X. S. Zhu, "Multi-labeled classification using maximum entropy method,in," *Proceeding sof SIGIR*, p. pp.274-281., 2005.
- [14] F. R. J. a. R. S. Kang, "Correlated label propagation with application to multi-label learning.," in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on. Vol. 2. IEEE*, 2006.
- [15] G. e. a. Chen, "Semi-supervised multi-label learning by solving a sylvester equation.," in *Proceedings of the 2008 SIAM International Conference on Data Mining. Society for Industrial and Applied Mathematics*, 2008.
- [16] T. J. Z. X.-S. Z.-J. Zha, "Graph-based semi-supervised learning with multiple labels.," *J. Vis. Commun. Image Represent.* 20, p. 97-103, 2009.
- [17] T. M. Mitchell, *Machine learning*. WCB, 1997.
- [18] K. B. N. d. F. a. D. F. P. Duygulu, "Object recognition as machine translation: Learning a lexicon for a fixed image vocabulary.pages," in *In Proceedings of the 7th European Conference on Computer Vision*, Copenhagen, Denmark., 2002.
- [19] X. M. K. N. a. Z.-H. Z. Kong, ". "Transductive multilabel learning via label set propagation.," *IEEE Transactions on Knowledge and Data Engineering* 25.3, pp. 704-719., 2013.
- [20] F. Sebastiani, "Machine learning in automated text categorization.," *ACM Compu. Surv.* 34 (1), p. 1-47., 2002.
- [21] G. e. a. Madjarov, "An extensive experimental comparison of methods for multi-label learning.," *Pattern Recognition* 45.9, pp. 3084-3104., 2012.
- [22] S. a. J. A. S. Mehrkanon, "Multi-label semi-supervised learning using regularized kernel spectral clustering.,"

سایر روش‌های یادگیری چندبرچسبی دقت پایین‌تری دارند. نکته دیگر اینکه رسیدن به این دقت، سرعت اجرا یا معیارها و پارامترهای دیگر را تحت‌الشعاع قرار نداده و در تمام معیارها به نسبت شاهد بهبود و یا هم تراز می‌مقادیر بوده‌ایم. دقت حاصل نتیجه اجرای چندین بار اجرا بر روی مجموعه داده‌های ذکر شده در مقاله و با انتخاب‌های متفاوت از مجموعه دادگان است که شائبه‌ی تأثیر نویز یا موارد دیگر را نفی می‌کند.

نکته دیگر اینکه نوع کار ما به‌مانند غالب کارهای نیمه نظارتی مبتنی بر گراف، ترالقایبی بوده و با روش‌های استقرایی متفاوت است؛ یعنی هدف آن برچسب زدن تصاویری است که خود آن‌ها در فرآیند آموزش دخالت دارند. بر این اساس دقت بهتر روش ما نسبت به روش ترالقایبی DLP اهمیت بسزایی دارد. همچنین روش ما در مقایسه با روش استقرایی mehrkanon در اکثر معیارها هم‌تراز و در مجموعه داده scene بسیار بهتر عمل نموده است. آنچه می‌توان در این حوزه به‌عنوان پژوهش‌های آینده مورد بررسی قرارداد استفاده از معیارهای شباهت کاراتر و الگوریتم‌های بهتر بجای روش KNN و همچنین بررسی الگوریتم‌های یادگیری نیمه نظارتی بر روی دیتا ست‌های چندبرچسبی و ارائه روش‌های جدید در این حوزه هست.

۷ سپاس‌گزاری

در اینجا بر خود لازم می‌دانیم از اساتید گران‌قدر آقای دکتر منصور جم‌زاد، استاد دانشگاه صنعتی شریف و دکتر سیدحمید امیری، استاد دانشگاه تربیت دبیر شهید رجائی که در خصوص روش‌های یادگیری نیمه نظارتی و مجموعه دادگان مورد استفاده، راهنمایی‌های ارزشمندی ارائه نمودند سپاس‌گزاری نماییم.

۸ مراجع

- [1] M.-L. a. Z.-H. Z. Zhang, "A review on multi-label learning algorithms.," *IEEE transactions on knowledge and data engineering* 26.8, pp. 1819-1837., 2014.
- [2] X. Zhu, "Semi-supervised learning literature survey," 2005.
- [3] B. a. J. T. Wang, "Dynamic label propagation for semi-supervised multi-class multi-label classification.," *Pattern Recognition* 52, pp. 75-84., 2016.
- [4] S. H. a. M. J. Amiri, "Automatic image annotation using semi-supervised generative modeling." , *Pattern Recognition* 48.1, pp. 174-188, 2015.
- [5] M. e. a. Zhao, "Automatic image annotation via compact graph based semi-supervised learning.," *Knowledge-Based Systems* 76, pp. 148-165., 2015.
- [6] O. B. T. L. J. W. B. S. D. Zhou, "Learning with local and global consistency," *Advances in Neural Information Processing Systems (NIPS)*, vol. 16, MIT Press, Cambridge, MA, 2004.

Neural Networks (IJCNN), 2016 International Joint Conference on. IEEE., 2016.

- [23] M. D. Smucker, J. Allan and B. Carterette, "A comparison of statistical significance tests for information retrieval evaluation," in *Proceedings of the sixteenth ACM conference on Conference on information and knowledge management*, 2007.



حسن ختن لو کارشناسی و کارشناسی ارشد را در رشته مهندسی کامپیوتر از دانشگاه علم و صنعت در سالهای ۱۳۷۴ و ۱۳۷۶ اخذ نموده و مقطع دکتری تخصصی را در رشته بینایی ماشین از دانشگاه پیر و ماری کوری پاریس فارغ‌التحصیل شده‌اند. در حال حاضر با رتبه دانشیاری عضو هیات علمی دانشگاه بوعلی سینا بوده و در موضوعات پردازش تصاویر پزشکی، بینایی ربات‌های اجتماعی، پردازش تصاویر و ویدئو مشغول به پژوهش می‌باشند.



محرّم منصوری‌زاده کارشناسی خود را در رشته مهندسی کامپیوتر در سال ۱۳۸۰ از دانشگاه اصفهان و کارشناسی ارشد و دکتری خود را در همین رشته در سالهای ۱۳۸۲ و ۱۳۸۸ از دانشگاه تربیت مدرس اخذ نموده است. او استادیار گروه مهندسی کامپیوتر دانشگاه بوعلی سینا است و به پژوهش در زمینه‌های یادگیری ماشین، ارتباط هوشمند انسان و ماشین، پردازش تصویر، ویدئو و صدا و همچنین زبان‌شناسی رایانشی مشغول است.



مجتبی کردآبادی کاردانی خود را در رشته نرم افزار در سال ۱۳۸۴ از دانشکده فنی شهید جباریان همدان و کارشناسی خود را در رشته مهندسی تکنولوژی نرم افزار در سال ۱۳۸۸ از دانشگاه جامع علمی کاربردی همدان اخذ نموده است و فارغ‌التحصیل کارشناسی ارشد مهندسی کامپیوتر، گرایش هوش مصنوعی در بهمن ماه سال ۹۶ از دانشگاه بوعلی سینا همدان است. او دبیر کامپیوتر در آموزش و پرورش و مدرس دروس رشته کامپیوتر در دانشگاه فنی و حرفه‌ای همدان می‌باشد و به پژوهش در حوزه یادگیری ماشین و داده‌کاوی و برنامه‌نویسی وب مشغول است.