

دریافت مقاله: ۹۳/۱/۲۷

پذیرش مقاله: ۹۳/۶/۲۰

## حاشیه‌نویسی تصویر با استفاده از الگوریتم خوشه‌بندی نیمه‌نظارتی طیفی

نگار صادق‌زاده،<sup>۱\*</sup> محبوبه شمسی،<sup>۲</sup> عبدالرضا رسولی کناری<sup>۳</sup>

---

<sup>۱</sup> دانشجوی کارشناسی ارشد هوش مصنوعی، دانشگاه آزاد اسلامی واحد علوم تحقیقات، کرمان، ایران

n.sadeghzade@gmail.com

<sup>۲</sup> استادیار، دانشکده فنی و مهندسی، دانشگاه صنعتی قم، قم، ایران

Shamsi@qut.ac.ir

<sup>۳</sup> استادیار، دانشکده فنی و مهندسی، دانشگاه صنعتی قم، قم، ایران

Rasouli@qut.ac.ir

---

چکیده: با توجه به رشد تصاویر دیجیتال در دنیای امروز، نیاز به روش‌هایی کارا برای حاشیه‌نویسی تصاویر احساس می‌شود. به دلیل اهمیت بالای این روش در فهم تصاویر دیجیتال، امروزه روش‌های مختلفی برای حاشیه‌نویسی تصاویر پیشنهاد داده شده‌اند. در این مقاله، از روشی بر مبنای خوشه‌بندی نیمه‌نظارتی طیفی استفاده شده است. استفاده از خوشه‌بندی نیمه‌نظارت شده باعث غلبه بر مسئله همگرا شدن در مینیمم محلی می‌شود، زیرا در این روش از داده‌های برجسب خورده توسط انسان استفاده می‌شود. همچنین، در روش پیشنهادی برای کاهش فاصله معنایی از سه مرحله بازخورد مرتبط استفاده شده است. عملکرد روش پیشنهادی روی ۵۰۰۰ داده مجموعه داده Corel تست شده و نتایج به دست آمده نشان‌دهنده کارایی و دقت فراوان روش پیشنهادی در مقایسه با روش‌های خوشه‌بندی دیگر است.

واژه‌های کلیدی: حاشیه‌نویسی تصاویر، یادگیری نیمه‌نظارتی، خوشه‌بندی طیفی نیمه‌نظارتی، بازخورد مرتبط.



## ۱. مقدمه

می‌شوند. مجموعه داده استفاده شده در این تحقیق، مجموعه داده Corel 5k است که شامل ۵۰۰۰ تصویر مختلف است. نتایج به دست آمده نشان‌دهنده کارایی بالای الگوریتم پیشنهادی نسبت به روش‌های موجود است.

در ادامه، در بخش دوم به مروری بر کارهای انجام شده در این حوزه می‌پردازیم. بخش سوم به استخراج ویژگی می‌پردازد. خوشه‌بندی نیمه‌نظارت شده در بخش چهارم توصیف می‌شود. بخش پنجم الگوریتم پیشنهادی را تشریح می‌کند. در بخش ششم، مجموعه داده مورد استفاده و آزمایش‌ها تشریح می‌شود و در نهایت، بخش هفتم به نتیجه‌گیری و کارهای آینده می‌پردازد.

## ۲. مروری بر ادبیات گذشته

بازیابی تصاویر به دلیل رشد سریع استفاده از تصاویر دیجیتال، یکی از زمینه‌های فعال تحقیق از سال ۱۹۹۰ بوده است. روش‌های متفاوتی برای حل مسئله بیان شده‌اند. از روش‌های دستی که توسط انسان انجام می‌شدند و بسیار هزینه‌بر، زمان‌گیر و محدود به دانش هر فرد در مورد تصاویر بودند [۱] گرفته تا روش‌های خودکار که به وسیله ماشین و بدون نظارت انجام می‌شدند. روش‌های خودکار با استفاده از تکنیک‌های مختلفی پیاده‌سازی شده‌اند؛ مانند روش‌های مبتنی بر تصویر که براساس ویژگی‌های استخراج شده از تصویر و با استفاده از روش‌های یادگیری ماشین دسته‌بندی می‌شوند [۲] یا روش‌های مبتنی بر متن که از تکنیک‌های متن‌کاوی و وب‌کاوی برای استخراج مفاهیم استفاده می‌کنند [۲ و ۳]. در واقع، حاشیه‌نویسی تصاویر به بازیابی و دسته‌بندی تصاویر گفته می‌شود که هر دسته بیانگر کلمه یا عبارتی در مورد تصویر است [۲]. در دنیای واقعی، دسته‌بندی تصاویر تنها براساس ویژگی‌های سطح پایین تصویر به چالشی به نام فاصله معنایی می‌انجامد. این فاصله تفاوتی است که بین آنچه ماشین به عنوان خروجی برمی‌گرداند و آنچه انسان به عنوان معنی برای تصویر در نظر می‌گیرد، به وجود می‌آید. برای کم کردن این فاصله سیستم‌های خودکار به سمت سیستم‌های نیمه‌خودکار

امروزه حاشیه‌نویسی تصاویر با توجه به رشد روزافزون تصاویر دیجیتال و نیاز به مدیریت و بازیابی کارای این تصاویر، به یکی از زمینه‌های پویا در تحقیق تبدیل شده است. هدف از حاشیه‌نویسی تصویر این است که کلماتی که بیانگر معنا و مفهوم واقعی تصاویر هستند، با تصویر همراه شوند. از آنجا که تفسیر این حجم از تصاویر توسط انسان بسیار دشوار، پرهزینه و زمان‌بر است، خودکار کردن حاشیه‌نویسی امری ضروری است. اما اطلاعات و ویژگی‌هایی که از تصاویر استخراج می‌شوند، همیشه بازگویی درست مفهوم تصویر نیستند و وجود فاصله معنایی میان معنا و مفهوم واقعی تصویر و آنچه سیستم به عنوان مفهوم تصویر استخراج می‌کند چالش اصلی سیستم‌های اتومات بدون ناظر هستند.

در سال‌های اخیر، تحقیقات در این زمینه به سمت سیستم‌های نیمه‌نظارت شده رفته است تا با استفاده از داده‌های کمکی که توسط کاربران تولید شده‌اند، این فاصله را کمتر کنند. روش‌های بسیاری در این زمینه ارائه شده‌اند. یکی از روش‌های پر استفاده در این زمینه، خوشه‌بندی است. خوشه‌بندی یک روش دسته‌بندی بدون ناظر است که تصاویری را که دارای ویژگی‌های بصری مشابهی هستند، در یک دسته قرار می‌دهد. اما دو مشکل اصلی بعضی الگوریتم‌های خوشه‌بندی، مقداردهی اولیه به مراکز خوشه و تأثیر معیار شباهت در کارایی آن است. در خوشه‌بندی اگر مراکز اولیه به درستی انتخاب نشود، ممکن است الگوریتم به مینیمم محلی همگرا شود.

در این مقاله، با استفاده از الگوریتم نیمه‌نظارتی طیفی و با استفاده از داده‌های برچسب‌دار، علاوه بر داده‌های بدون برچسب در انتخاب مراکز اولیه خوشه‌ها بر مشکل همگرا شدن در مینیمم محلی فائق آمده و برای کمتر شدن فاصله معنایی از سه مرحله بازخورد مرتبط استفاده شده است. در این مراحل، نظریات کاربران درباره تصاویری که نامربوط به خوشه‌اند یا تصاویری که مربوط به خوشه‌اند، گردآوری می‌شود. همچنین، مراکز دقیق خوشه‌ها با دقت بالا تعیین

است، به بازخوردهای فراوانی نیاز دارد و از سرعت پایینی برخوردار است. سیستم معرفی شده توسط چنگ چین چانگ با نام IGAnn از نظر کاربران برای برجسب‌های متنی تصاویر استفاده می‌کند [۸]. در سیستم پیشنهاد داده شده، تصاویر مرتبط یا نامرتبط با یک برجسب متنی و با نظر کاربر مشخص می‌شوند و سپس، طی تکرارهایی با استفاده از یک دسته‌بندی‌کننده، ارزش اطمینان برای تصاویر برجسب نخورده محاسبه می‌شوند. در [۹] یک ساختار شاخص جدید همراه با یک روش پردازش پرس‌وجو پیشنهاد داده شده است که تصویر هدف را با تکرارهای کمتری به دست می‌آورد. همچنین با استفاده از روش‌هایی، تأثیر بازخوردهای نادرست کاربر را به حداقل می‌رساند. در [۱۰] از خوشه‌بندی سلسله مراتبی و K-میانگین رای بازیابی محتوی تصویر استفاده شده است. این سیستم از ویژگی‌های رنگ استخراج شده از تصویر، برای گروه‌بندی تصاویر درون خوشه‌ها استفاده می‌کند.

### ۳. استخراج ویژگی

همان‌طور که ما با نگاه کردن به یک تصویر می‌توانیم آن را تفسیر کرده و شرح دهیم، باید دید آیا ماشین نیز می‌تواند این کار را انجام دهد؟ در حاشیه‌نویسی تصاویر، هدف این است که از مشخصه‌های بصری یک عکس به محتوای آن پی ببریم. ویژگی‌های ظاهری یا بصری هر تصویر، در قالب ویژگی‌های سراسری و محلی مطرح می‌شوند. ویژگی‌هایی نظیر رنگ، بافت و شکل تصویر اصلی‌ترین ویژگی‌ها برای استخراج از تصویرند. البته پیدا کردن ویژگی‌های مؤثر برای بازیابی تصاویر و حاشیه‌نویسی آن‌ها در خوشه‌بندی تصاویر بسیار مهم است و در کارایی سیستم‌های حاشیه‌نویسی تصویر تأثیر مستقیم دارد. روش‌های مختلفی برای استخراج ویژگی‌های یک تصویر وجود دارد و همچنین، ویژگی‌های مختلفی را بسته به کاربرد می‌توان از تصاویر مورد نظر استخراج کرد. در ادامه، بعضی از روش‌های استفاده شده در این مقاله، برای به دست آوردن ویژگی‌های مخصوصی از تصویر آورده شده‌اند.

متماثل شده‌اند که با تعامل با کاربر، کارایی سیستم بهبود قابل توجهی پیدا می‌کند. این تعامل تحت عنوان بازخورد مرتبط<sup>۱</sup> بین کاربر و سیستم، در دهه ۱۹۹۰ در سیستم‌های بازیابی تصویر بر مبنای محتوا با قصد آوردن کاربر در حلقه بازیابی برای کاهش فاصله معنایی بین اینکه پرس‌وجو چه می‌گوید و کاربر چه فکر می‌کند، معرفی شده است. مقایسه میان روش‌های مختلف استخراجی که در سیستم‌های مشابه به کار گرفته شده‌اند، با توجه به بازه وسیع روش‌های پیاده‌سازی شده برای نشان دادن شباهت و وابستگی مجموعه تصاویر، یک مسئله چالش برانگیز است. چین و همکارانش، سیستمی طراحی کردند که از خوشه‌بندی K-میانگین با محدودیت‌های pair-wise برای حاشیه‌نویسی استفاده می‌کرد [۴]. فرض اساسی این روش این بود که حاشیه‌نویسی‌هایی که بسیار مرتبط بودند، نگهداری می‌شدند و مابقی حذف می‌شدند. به این ترتیب، تنها اطلاعات متنی کلی استفاده می‌شد و پردازش مستقل از تصویر هدف صورت می‌گرفت و تصاویر متفاوتی که حاشیه‌های کاندید یکسان داشتند نتایج حاشیه‌نویسی آن‌ها یکسان می‌شد. سریکانس و همکارانش روشی را ارائه کردند که در آن، از یک دسته‌بندی‌کننده سلسله‌مراتبی برای استخراج برجسب‌ها از آنتولوژی متنی استفاده می‌شود [۵]. در سال‌های اخیر، توجه به روش‌های خوشه‌بندی در حال افزایش است. در [۶] ابتدا از تئوری گراف و روش NCut برای خوشه‌بندی استفاده کرده است، سپس از بازخورد مرتبط برای تصحیح معیار شباهت بر اساس خواسته کاربر بهره برده است. این روش علی‌رغم قابلیتی که دارد نمی‌تواند نگاشت مستقیمی بین ویژگی‌های سطح پایین و مفاهیم سطح بالا ایجاد کند. همچنین زمان محاسباتی زیادی نیاز دارد. [۷] از خوشه‌بندی K-میانگین نیمه‌نظارت شده برای حاشیه‌نویسی تصاویر پزشکی استفاده کرده است و بازخورد مرتبط را برای بهبود کارایی به کار برده است. این روش به دلیل اینکه تنها از بازخورد مرتبط برای خوشه‌بندی تصاویر استفاده کرده و از هیچ روش خاصی برای یادگیری نظارت کاربر بهره نبرده

1. relevance feedback

$$\text{even Image} = \frac{(\text{Original Image} + \text{Flip})}{2} \quad (1)$$

متوسط سطری ستونی، متوسط قطری، متوسط ستونی ردیفی می‌توانند به‌عنوان بردار ویژگی در نظر گرفته شوند.

### ۳.۳. استخراج ویژگی هیستوگرام بافت الگوی باینری محلی<sup>۱</sup>

تکنیک الگوی باینری محلی یکی از روش‌های طبقه‌بندی بافت تصویر است [۱۴]. در این روش، ابتدا با در نظر گرفتن یک همسایگی از تصویر، شدت روشنایی نقاط موجود در این همسایگی با شدت روشنایی مرکز همسایگی مقایسه می‌شود. معمولاً برای اینکه این عملگر نسبت به چرخش تصویر حساس نباشد، همسایگی را به‌صورت حلقوی در نظر می‌گیرند و مختصات نقاطی را که دقیقاً روی مرکز پیکسل قرار ندارند، با درون‌یابی پیدا می‌کنند. روش LBP در هر همسایگی از تصویر با رابطه (۲) تعریف می‌شود [۱۵].

$$LBP_{P,R} = \sum_{i=0}^{P-1} s(g_i - g_c) 2^i \quad (2)$$

به‌طوری‌که P تعداد نقاط همسایگی مرکز،  $g_i$  شدت روشنایی نقاط همسایگی مرکز،  $g_c$  شدت روشنایی نقطه مرکزی و  $s(\cdot)$  تابع علامت است که به‌صورت رابطه (۳) تعریف می‌شود:

$$s(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases} \quad (3)$$

به‌منظور تولید بردار ویژگی تصویر، می‌توان از هیستوگرام استفاده کرد. این هیستوگرام P+2 بازه دارد که فراوانی رخداد نقاط با برچسب‌های مختلف در تصویر را نشان می‌دهد.

### ۴.۳. تشخیص ویژگی‌های SIFT

یکی از روش‌های تشخیص و توصیف ویژگی‌های محلی تصویر در بینایی ماشین، روشی موسوم به SIFT<sup>۲</sup> است. این الگوریتم به‌وسیله دیوید لو<sup>۳</sup> در سال ۱۹۹۹ ابداع شد [۱۶]. به‌کمک این روش، ویژگی‌هایی از تصویر استخراج می‌شوند

### ۱.۳. استخراج ویژگی تشخیص لبه

در متد تشخیص لبه، تصویر باینری اسکن می‌شود تا لبه‌ها پیدا شوند. یکی از روش‌هایی که می‌توان بر پایه آن لبه‌های تصویر را تشخیص داد اسکن تصویر بر پایه متد K نزدیک‌ترین همسایه است. یک پیکسل از پیش‌زمینه مانند P و مجموعه‌ای از پیکسل‌های پیش‌زمینه متصل به P که آن‌ها را اجزای دربرگیرنده P می‌نامیم، در نظر می‌گیریم [۱۱]. ردیابی با تنظیم موقعیت اولیه برای مشخص کردن مبدأ لبه انجام می‌شود. بردار ویژگی که از این روش محاسبه می‌شود، توصیف فوریه نامیده می‌شو و با کمک توصیف فوریه، ضرایب فوریه محاسبه می‌شوند. برای اطمینان از اینکه لبه‌ها به‌طور کامل پیدا شده‌اند، باید موقعیت ابتدا و انتها به‌هم‌رسیده باشند و در واقع برابر باشند.

در اینجا از روشی مبتنی بر شکل و بدون بخش‌بندی تصویر برای پیدا کردن لبه‌های تصویر استفاده شده است. الگوریتم استفاده‌شده برای استخراج ویژگی لبه تصویر، الگوریتم هیستوگرام جهت لبه MPEG-7 است. هیستوگرام جهت لبه نحوه گسترش لبه‌ها در تصویر را نشان می‌دهد. در این روش، با دسته‌بندی لبه‌های تصویر روی جهت لبه، هیستوگرامی از فراوانی لبه‌ها ایجاد می‌شود و از آن به‌عنوان بردار ویژگی تصویر استفاده می‌شود. برای مشاهده جزئیات مربوط به هیستوگرام جهت لبه به [۱۲] مراجعه شود.

### ۲.۳. استخراج ویژگی متوسط رنگ

تکنیک‌های متوسط رنگ یکی از روش‌های استخراج ویژگی در دامنه فضایی است. استخراج ویژگی در دامنه فضایی، به‌دلیل کاهش اندازه بردار ویژگی انجام می‌شود. این تکنیک از تصویر زوج استفاده می‌کند. برای ساختن تصویر زوج از تصویر اصلی، ابتدا باید تصویر فلیپ آن ساخته شود. تصویر فلیپ یک کپی از تصویر اصلی است که در راستای بردار X و Y دوران داده شده است. سپس از مجموع تصویر اصلی و تصویر فلیپ، تصویر زوج به‌دست می‌آید [۱۳]. به این ترتیب، تصویر زوج از رابطه زیر به‌دست می‌آید:

1. Local Binary Pattern(LBP) texture histogram  
2. Scale Invariant Feature Transform  
3. David Lowe

حوزه مکان تعریف شده اند و برای استخراج محتوای فرکانسی وابسته به راستا مناسب اند [۲۰ و ۲۱]. فیلترهای گابور که در پردازش تصویر به کار می‌روند، دو بعدی هستند و هر فیلتر دارای فرکانس و جهت مشخصی است. این فیلترها به‌عنوان ابزاری برای استخراج ویژگی‌های تصاویر به کار برده می‌شوند و مستقیماً روی تصویر دیجیتال، در حوزه مکان اعمال می‌شوند. فیلتر گابور  $G(x,y)$ ، حاصل مدولاسیون سیگنال‌های سینوسی  $s(x,y)$  و گاوسی  $g(x,y)$  است.

$$G(x,y)=g(x,y)s(x,y) \quad (5)$$

سیگنال سینوسی فیلتر گابور به فرم عبارت ۶ است.

$$s(x,y)=\exp(j2\pi(ux+vy)) \quad (6)$$

در عبارت بالا،  $x$  و  $y$  مختصات نقاط در تصویرند.  $u$  و  $v$  هم فرکانس‌های مرکزی سیگنال سینوسی در راستای محور  $x$  و  $y$  هستند. می‌توان فرکانس مکانی سیگنال سینوسی را که در مختصات کارتزین تعریف می‌شود، به صورت قطبی نیز بیان کرد.

$$F=\sqrt{u^2+v^2}, \theta=\tan^{-1}\left(\frac{v}{u}\right) \quad (7)$$

با استفاده از عبارت بالا رابطه سیگنال سینوسی را می‌توان به صورت عبارت (۸) نیز به کار برد.

$$s(x,y)=\exp(j2\pi F(x\cos\theta+y\sin\theta)) \quad (8)$$

که  $F$  بیانگر فرکانس و  $\theta$  جهت فیلتر است. پوشش گاوسی<sup>۲</sup> فیلتر گابور به فرم عبارت (۹) است.

$$G(x,y)=\exp\left(-\frac{1}{2}\left[\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right]\right) \quad (9)$$

$\sigma_x$  و  $\sigma_y$  به ترتیب، انحراف معیار تابع پوشش گاوسی در راستای محورهای  $x$  و  $y$  هستند. اگر  $\sigma_x = \sigma_y$  در نظر بگیریم، فیلتر گابور با عبارت شماره ۱۰ نشان داده می‌شود.

$$G_{F,\theta,\sigma}(x,y) = \exp\left\{-\frac{x^2+y^2}{2\sigma^2}\right\} * \exp\{j2\pi F(x\cos\theta + y\sin\theta)\} \quad (10)$$

که نسبت به تغییر مقیاس و چرخش مقاوم‌اند. به عبارت دیگر، اگر ویژگی SIFT یک شیء یا الگو در تصویری شناسایی شود، به طوری که این شیء یا الگو در تصویر دیگری با مقیاس یا موقعیت متفاوتی ظاهر شود، ویژگی‌های SIFT آن تغییر چندانی نخواهد داشت. این امر کمک فراوانی به شناسایی اشیا در تصاویر مختلف می‌کند. این ویژگی‌ها نسبت به تغییرات جزئی نور و نویز نیز مقاوم‌اند. از دیگر خاصیت‌های این روش، سرعت استخراج ویژگی‌ها از تصاویر است [۱۷].

### ۵.۳. استخراج ویژگی با استفاده از ماتریس وقوع

ماتریس وقوع روشی آماری برای بررسی بافت تصویر است. هر عنصر از این ماتریس، آماری از وقوع نقاطی با مقادیر و ارتباط مکانی مشخص است. به بیان دقیق‌تر، ماتریس وقوع یک ماتریس مربعی است که اندازه آن به بیشینه شدت خاکستری نقاط در تصویر بستگی دارد و درایه  $P(i,j)$  نشان‌دهنده تعداد وقوع نقطه‌ای با شدت خاکستری  $i$  با فاصله مشخص  $d$  از نقطه‌ای با شدت خاکستری  $j$  است. علاوه بر فاصله، زاویه بین نقاط را نیز می‌توان تعیین کرد. زوایای بین دو نقطه می‌توانند با صفر، ۴۵، ۹۰ و ۱۳۵ بیان شوند. بعد از انتخاب فاصله و جهت مناسب و تشکیل ماتریس وقوع، باید به دنبال استخراج ویژگی‌های مناسبی بود که متمایزکننده یک کلاس از دیگری باشند [۱۸ و ۱۹]. یکی از ویژگی‌های قابل استخراج از ماتریس وقوع کنتراست است. ویژگی کنتراست طبق عبارت ۴ محاسبه می‌شود.  $P(i,j)$  درایه‌های ماتریس وقوع و  $i$  و  $j$  مقادیر شدت هستند.

$$\text{contrast} = \sum_{i,j} |i - j|^2 p(i,j) \quad (4)$$

### ۶.۳. استخراج ویژگی با استفاده از فیلتر گابور

فیلترهای دوبعدی گابور فیلترهای میان‌گذر حساس به فرکانس و راستا هستند که به طور بهینه در حوزه فرکانس و

ماتریس به صورت  $W = \{w_{ij}\}$  و  $\bar{W} = \{\bar{w}_{ij}\}$  استفاده می‌شود که به ترتیب، مشخص‌کننده وزن قیود Must-link و Can-not-link است. در سال ۲۰۰۲، کلین طی مقاله‌ای ادعا کرد که مجبور کردن توابع یادگیری به ارضا نمودن قیود مسئله نمی‌تواند همیشه نتایج خوبی داشته باشد [۲۲]. او در ادامه اعلام کرد که داده‌های برچسب‌داری که کاربر در اختیار الگوریتم قرار می‌دهد، علاوه بر اینکه راجع به خود داده‌ها اطلاعات می‌دهند، در مورد همسایگان خود نیز اطلاعاتی در خود نهفته دارند. او اطلاعات درباره چگونگی خود داده را اطلاعات سطح نمونه<sup>۱</sup> و اطلاعات همسایگی آن را اطلاعات سطح فضا<sup>۲</sup> نامید. در ادامه، او بیان کرد که در الگوریتم‌های دیگر نظیر K-میانگین نیمه‌نظارتی، در صورتی که اطلاعاتی درباره داده‌ای وجود داشته باشد، این اطلاعات تنها در خصوص همان داده مورد استفاده قرار می‌گیرد. در صورتی که فرض شود که فرد دهنده اطلاعات می‌خواهد که داده‌های پرت مجدداً مورد خوشه‌بندی قرار گیرد، این روش‌ها مناسب‌اند. اما در بیشتر مواقع، داده‌های برچسب‌دار مشخص‌کننده خود و خوشه خود هستند. کلین در این مقاله سعی کرد که داده‌های برچسب‌دار را به صورتی در الگوریتم خود وارد کند که میزان فاصله داده‌های مجاور، نمایشی از برچسب‌های داده‌های برچسب‌دار باشد. این الگوریتم بر پایه خوشه‌بندی تراکمی سلسله‌مراتبی پایه‌گذاری شد. هدف از این الگوریتم، آن است که ابتدا میزان شباهت داده‌ها در فضای ویژگی محاسبه گردد، سپس با در اختیار داشتن مجموعه کوچکی از داده‌های برچسب‌دار یا اطلاعات اضافی داده‌ها را در فضایی شبیه به فضای ویژگی به گونه‌ای تغییر داده شود که قیود به میزان زیادی برآورده شود. این تغییرات ممکن است

یکی از مزایای فیلتر گابور این است که می‌تواند روی جهت‌های مختلف در تصاویر کار کند و در نتیجه، ویژگی‌ها را روی جهت‌های مختلف استخراج کند. به منظور استفاده از فیلترهای گابور در استخراج ویژگی، معمولاً بانکی از فیلترها تهیه می‌شود و پاسخ تصاویر به فیلترها مورد بررسی قرار می‌گیرد. نکته مهم در تولید بانک فیلترها چگونگی انتخاب پارامترهای فرکانس و جهت و انحراف معیار هر کدام از فیلترهاست. می‌توان این پارامترها را به طور تجربی و با اطلاع از محتوای تصاویر انتخاب کرد.

#### ۴. خوشه‌بندی نیمه‌نظارتی

خوشه‌بندی در تحلیل داده‌ها، یادگیری ماشین و شناسایی الگو یکی از بخش‌های اساسی و جدا نشدنی به‌شمار می‌رود. به همین دلیل، بهبود نتایج حاصل از خوشه‌بندی از همان ابتدا به شدت مورد توجه محققان بوده و اهمیت بسزایی داشته است. یکی از روش‌های نوین برای بهبود کیفیت روش‌های مختلف در خوشه‌بندی، استفاده از یادگیری نیمه‌نظارتی است. همین‌طور که از نام آن پیداست، قلب این نوع الگوریتم‌ها یک الگوریتم خوشه‌بندی است و سعی بر آن است که با استفاده از تعداد محدودی داده‌های برچسب‌دار، نتایج بهتری به دست آورده شود. در برخی الگوریتم‌های نیمه‌نظارتی، از اطلاعاتی به صورت قیودهای Must-link و Can-not-link استفاده می‌شود. این نوع اطلاعات برای جمع‌آوری، بسیار ساده‌تر از برچسب‌های واقعی داده‌ها هستند. چون شخص خبره ممکن است نتواند گروه واقعی یک داده را مشخص کند، اما به راحتی می‌تواند مشخص کند که دو داده در یک گروه‌اند یا خیر. به این صورت که داده‌هایی را که باید در یک گروه قرار گیرند، در مجموعه  $M$  و داده‌هایی را که در یک گروه نیستند، در  $C$  قرار دهد. همچنین، برای مشخص کردن وزن هر قید از دو

1. Instance-level information  
2. Space-level information

اما یکی از بزرگ‌ترین تفاوت‌های این دو آن است که کار کلین بر پایه الگوریتم سلسله‌مراتبی تراکمی بنا نهاده شده بود، درحالی‌که چن از الگوریتم خوشه‌بندی طیفی کمک گرفته است. دیگر تفاوت اینکه زمانی که کلین داده‌های Must-link را با فاصله صفر در نظر می‌گیرد، در میان داده‌ها نقاطی به وجود می‌آید که برای آن نمی‌توان فاصله‌ای در نظر گرفت. اما چن در اینجا از نزدیک‌ترین همسایه استفاده کرده است. در این حالت از به وجود آمدن چنین نقاطی جلوگیری خواهد شد. علاوه بر آن، روش کلین داده‌های Can-not-link را در نظر نگرفته بود، اما در روش حاضر، این مورد نیز گنجانده شده است.

چن با ایده گرفتن از چگونگی ارتباط خانواده‌ها با یکدیگر مدلی ارائه کرد که به کمک آن بتواند داده‌های برجسب‌دار را به الگوریتم خوشه‌بندی طیفی اضافه کند. او نام مدل خود را غریبه‌های نزدیک یا خویشاوندان دور (NSDR)<sup>۴</sup> نامید. این مدل می‌گوید که اگر دو نفر دارای روابط خویشاوندی باشند، احتمال آنکه آن دو فرد خصوصیات نزدیکی داشته باشند، بسیار بالاست؛ برای مثال، اگر دو داده  $A_i$  و  $B_j$  با یکدیگر برادر باشند، اما داده  $C_j$  با آن دو روابطی نداشته باشد، آنگاه از میان بچه‌های این سه داده، انتظار می‌رود که بچه‌های داده‌های  $A_i$  و  $B_j$  شباهت بیشتری به هم داشته باشند. این موضوع بیان می‌نماید که داده‌هایی که با هم شبیه‌اند، باید به همراه همسایگان‌شان با یکدیگر فاصله کمی داشته باشند، درحالی‌که داده‌هایی که به هم شبیه نیستند، باید به همراه همسایگان‌شان از یکدیگر دور باشند. این حالت به صورت شکل ۱ قابل تعریف است. در قسمت (a)،  $A_{i0}$  و  $B_{j0}$  خویشانان دورند و به همین دلیل، خودشان و همسایگان‌شان را مقید به نزدیک شدن به هم می‌کنند. در قسمت (b)،  $A_{i0}$  و  $C_{j0}$  غریبه‌های نزدیک‌اند و خودشان و همسایگان‌شان را مقید به دور شدن از هم می‌کنند.

به صورت تغییرات بنیادی در فضای ویژگی خود را نشان دهد که می‌تواند باعث شناسایی خوشه‌های جدیدی شود یا با تغییر شکل‌هایی سطحی باعث بهتر شدن مرزهای خوشه‌ها شود. در این الگوریتم، سعی شده است که دو هدف کلی دنبال شود. اول آنکه دو داده که در یک خوشه قرار می‌گیرند، بسیار به هم نزدیک باشند که این کار به نام تحمیل قیود<sup>۱</sup> نام‌گذاری شده است؛ دوم اینکه داده‌های بدون برجسب دو اصل زیر را رعایت کنند:

■ فرض (۱-۴): اگر داده  $x_i$  و  $x_j$  به هم خیلی نزدیک‌اند، بنابراین داده‌هایی که خیلی به نقطه  $x_i$  نزدیک هستند، به داده  $x_j$  نزدیک باشند.

■ فرض (۲-۴): اگر داده  $x_i$  و  $x_j$  خیلی دور از هم‌اند، بنابراین داده‌هایی که خیلی به نقطه  $x_i$  نزدیک هستند، از داده  $x_j$  دور باشند.

به این کار انتشار قیود<sup>۲</sup> گویند. ذکر این نکته الزامی است که در صورتی که فضای مسئله تنها از داده‌های بدون برجسب تشکیل شده بود، این قیود خود به خود مرتفع می‌شد. اما به دلیل تغییر فضا این خصوصیت از بین می‌رود. در قلب این الگوریتم، الگوریتم خوشه‌بندی تراکمی وجود دارد که برای یافتن فاصله دو خوشه دو داده با بیشترین فاصله را برمی‌گزیند.

#### ۱.۴ خوشه‌بندی نیمه‌نظارتی طیفی<sup>۳</sup>

در سال ۲۰۱۲، چن اولین ایده را در زمینه خوشه‌بندی نیمه‌نظارتی طیفی ارائه کرد [۲۳]. اگرچه ایده اجرا شده در این مقاله نیز به همان سادگی الگوریتم‌های دیگر است، نتایجی که از این الگوریتم به دست آمده، بر نتایج آن‌ها برتری دارد. الگوریتم پیشنهادی چن، الهامی عمیق از الگوریتم کلین است؛

1. Imposing constraints
2. Propagating constraints
3. Semi-supervised spectral clustering



شباهتشان برابر بیشترین مقدار شباهت بین دو داده و داده‌های همسایه قرار داده می‌شود.

$$\begin{aligned} & \text{for}(v_{i_0}, v_{j_0}) \in M \\ & \text{for}\{v_i \in N(v_{i_0}) \text{ and } v_j \in N(v_{j_0})\} \\ & S_{ij} = f(v_i, v_j) \\ & = \max\{S_{ij}, S_{i i_0} \cdot S_{i_0 j_0} \cdot S_{j j_0}\} \end{aligned} \quad (11)$$

پس از پایان این مرحله به تحمیل داده‌های Can-not-link پرداخته می‌شود:

۱. در ماتریس شباهت در صورتی که داده‌های دارای برچسب Can-not-link باشند، شباهتشان برابر صفر در نظر گرفته می‌شود.

۲. برای داده‌هایی که با هم شبیه نیستند، باید داده‌های همسایشان با هم شباهت کمی داشته باشند. پس طبق فرمول رابطه (۱۲) میزان شباهتشان مشخص می‌شود. این فرمول می‌گوید که داده‌هایی که همسایه داده‌های ناشیبه هستند، اما همسایه دیگری نیستند، باید شباهتی حداقل نظیر داده‌های Can-not-link به صورت عبارت زیر داشته باشند.

$$\begin{aligned} & \text{for}(v_{i_0}, v_{j_0}) \in C \\ & \text{for}\{v_i \in N(v_{i_0}) \text{ and } v_j \in N(v_{j_0}) \text{ and } \\ & v_i \neq N(v_{i_0}) \text{ and } v_j \neq N(v_{i_0})\} \\ & S_{ij} = h(v_i, v_j) = \min\{S_{ij}, 1 - S_{i i_0} \cdot S_{j j_0}\} \end{aligned} \quad (12)$$

الگوریتم این روش به شکل زیر است:

قبل از اینکه چگونگی استفاده از این مدل در خوشه‌بندی طیفی توضیح داده شود، به معرفی تعدادی نماد نیاز است.

$G(V, E, S)$  گرافی بدون جهت را مشخص می‌کند که دارای نودهای  $V = \{v_1, v_2, \dots, v_n\}$  و لبه‌های  $E$  است. همچنین این گراف دارای ماتریس شباهت  $S$  است که میزان شباهت داده‌ها را مشخص می‌کند. مجموعه داده‌های مشخص در یک گروه را به صورت  $M = \{(v_i, v_j)\}$  و داده‌هایی که در یک گروه نیستند، به صورت  $C = \{(v_i, v_j)\}$  نمایش داده می‌شود. برای مجموعه‌های  $M$  و  $C$  و برای انتشار برچسب‌ها به صورت زیر عمل می‌شود:

۱. به علت اینکه قیود Must-link خاصیت تعدی دارند، باید داده‌های دارای خاصیت تعدی گسترش داده شود.  
۲. داده‌هایی که خاصیت Must-link دارند، به شدت به هم شبیه‌اند. پس در ماتریس شباهت، مقدار آن‌ها برابر ۱ در نظر گرفته می‌شود.

۳. جست‌وجو در جهت یافتن داده‌های بدون برچسبی که با داده‌های Must-link همسایه باشند، صورت می‌گیرد. میزان شباهت این داده‌ها به صورت عبارت زیر به روز درمی‌آید. بر طبق این فرمول، همسایگان دو داده که برچسب Must-link دارند، یافت می‌شود و

محدوده خوشه کافی نیست و ممکن است برخی از تصاویر اشتباهی در خوشه‌های همسایه قرار بگیرند به این علت که فاصله تصویر تا خوشه همسایه، کمتر از خوشه درست است. برای حل این مشکلات از بازخورد مرتبط استفاده می‌شود. در شکل ۲، فلوجارت سیستم پیشنهادی نشان داده شده است. همان‌طور که در فلوجارت دیده می‌شود، ابتدا ویژگی‌های سطح پایین از تصاویر، با استفاده از روش‌های گفته‌شده از پایگاه داده استخراج می‌شود و بردار ویژگی‌ها ساخته می‌شود، بردار ویژگی‌های ساخته شده همراه با اطلاعاتی از جانب کاربر که حاوی تعداد کمی تصاویر حاشیه‌نویسی شده از هر خوشه است، به‌عنوان ورودی به مرحله بعد فرستاده می‌شود. مراکز اولیه خوشه‌ها به‌صورت تصادفی از بین تصاویر حاشیه‌نویسی شده انتخاب می‌شوند. در واقع، هر خوشه تعدادی تصویر حاشیه‌نویسی شده صحیح دارد که توسط کاربر حاشیه‌نویسی شده‌اند و انتخاب مراکز خوشه‌ها از بین همین تصاویر انجام می‌شود. علاوه بر این، در ابتدا دو مجموعه C و M نیز از بین داده‌های برجسب‌دار تشکیل می‌شوند و به‌عنوان ورودی به الگوریتم NSDR فرستاده می‌شوند. سپس در مرحله بعد، تصاویر پایگاه داده توسط الگوریتم NSDR خوشه‌بندی می‌شوند.

بعد از خوشه‌بندی، ماتریس شباهت با توجه به بازخورد کاربر و توسط الگوریتم NSDR به‌روز می‌شود و دقت سیستم بهبود پیدا می‌کند. کاربر با تشخیص اینکه تصویر مربوط یا نامربوط به خوشه انتخاب شده است، به انتخاب‌های بعدی الگوریتم کمک خواهد کرد. بازخورد مرتبط تا سه مرحله انجام می‌شود.

## ۶. نتایج تجربی

### ۱.۶. مجموعه داده

یکی از مجموعه تصاویری که برای آموزش و آزمایش سیستم‌های حاشیه‌نویسی تصاویر استفاده می‌شود مجموعه Corel [۲۴] است. این مجموعه به یکی از مجموعه‌های معیار برای حاشیه‌نویسی تصاویر تبدیل شده است.

ورودی: یک مجموعه داده بدون چسب به‌صورت  $V = \{v_1, v_2, \dots, v_n\}$ ، تعداد خوشه‌ها  $k$ ، مجموعه‌ای از داده‌ها به‌صورت قیود Must-link و Can-not-link به شکل  $M = \{(x_i, x_j)\}$  و  $C = \{(x_i, x_j)\}$  پارامتر  $\sigma$ ، تعداد همسایگان  $k$ .

۱. از روی داده‌های  $V$  گراف  $G$  ساخته می‌شود.

۲. تابع فاصله‌ای به‌صورت  $d(v_i, v_j)$  به‌عنوان کوتاه‌ترین مسیر که دو داده  $v_i$  و  $v_j$  تعریف می‌شود.

۳. ماتریس شباهت به‌صورت عبارت  $S = (s_{ij} = e^{-\frac{d^2(v_i, v_j)}{2\sigma^2}})$  محاسبه می‌شود.

۴. داده‌های برجسب‌دار طبق الگوی عبارات ۱۱ و ۱۲ تحمیل کرده و انتشار داده می‌شود.

۵. ماتریس‌های  $W = D^{-1}S$  و  $D = \text{Diag}(S\mathbf{1})$  محاسبه می‌گردد.

۶. بزرگ‌ترین  $k$  بردار ویژگی از ماتریس  $W$  محاسبه می‌شود و ماتریس  $Y$  به‌صورت  $Y = [y_1, y_2, \dots, y_n]$  تشکیل داده می‌شود.

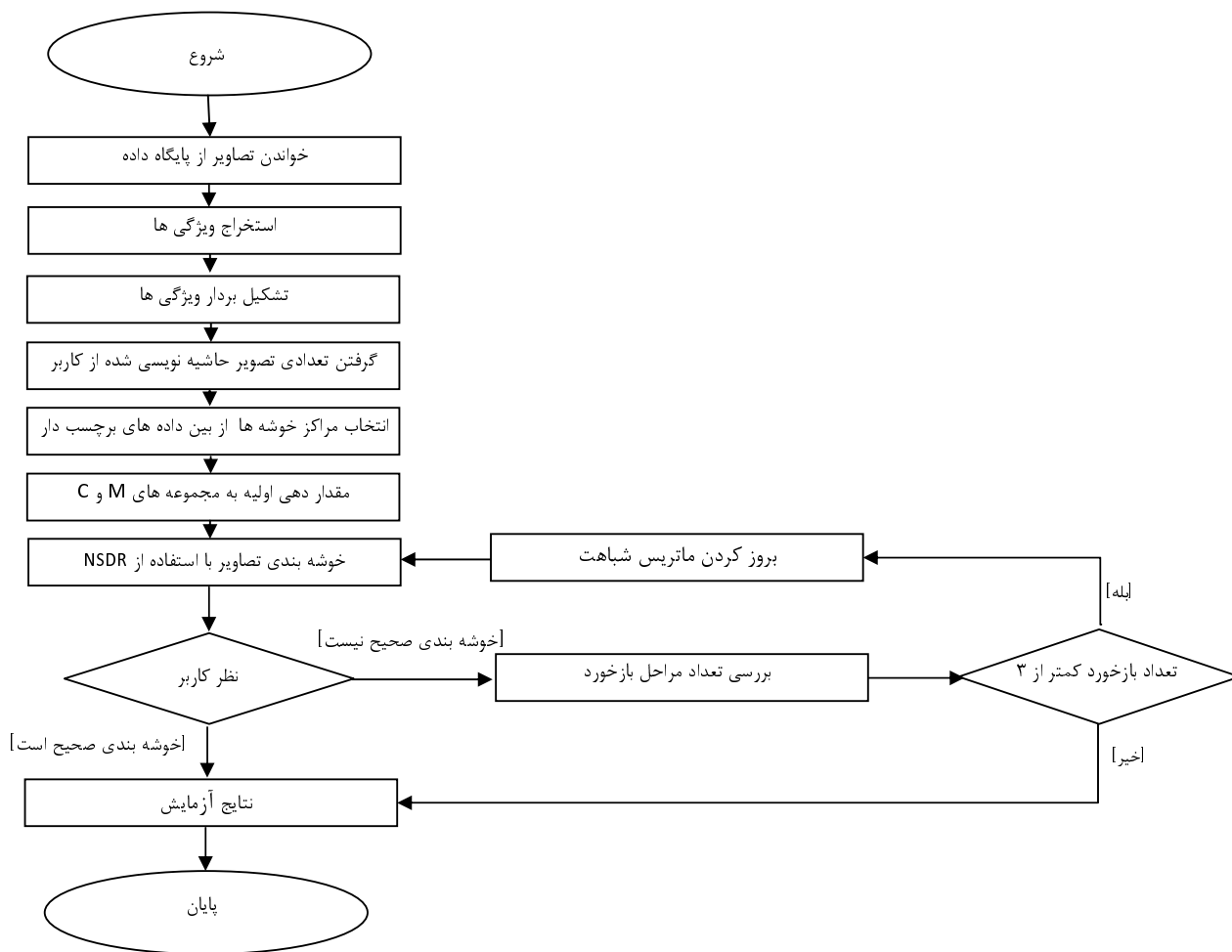
۷. هر سطر از ماتریس  $Y$  به‌صورت یک نقطه در فضای  $k$  بعدی در نظر گرفته می‌شود. سپس توسط الگوریتم K-میانگین خوشه‌بندی انجام می‌شود.

۸. به‌ازای هر سطر  $z$  در ماتریس  $Y$  به خوشه  $i$  اختصاص یافته، داده  $z$  به خوشه  $i$  اختصاص داده می‌شود.

خروجی: تعداد  $k$  خوشه جدا از هم.

## ۵. الگوریتم پیشنهادی

در الگوریتم پیشنهادی، از خوشه‌بندی نیمه‌نظارت‌شده تحت الگوریتم NSDR معرفی شده در [۲۲] و سه مرحله بازخورد مرتبط استفاده شده است. از آنجایی که تنوع در داده‌های مورد بررسی زیاد است، صرفاً استفاده از ویژگی‌های سطح پایین نمی‌تواند منجر به خوشه‌بندی دقیقی شود؛ بنابراین، پیدا کردن

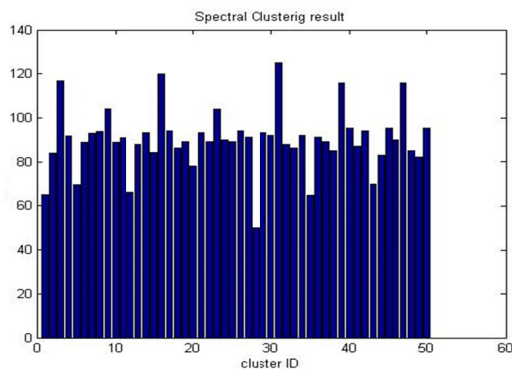


شکل (۲): فلوچارت الگوریتم پیشنهادی

تصویر از ۳.۵ کلمه برای حاشیه‌نویسی استفاده شده است. همان‌طور که ذکر شد، تعداد کل کلمات به‌کاررفته در حاشیه‌نویسی تصاویر مجموعه آموزش تعیین‌کنند تعداد کل دسته‌هاست. در واقع، به هر کلمه شماره‌ای متناظر با دسته تخصیص داده می‌شود؛ برای مثال، اگر شماره دسته  $k$  به کلمه «هوایما» نسبت داده شود، بدین معناست که تصویری که کلمه «هوایما» در آن حاشیه‌نویسی شده عضو دسته یا کلاس  $K$  است. همان‌طور که در بخش قبل گفته شد، برای ارزیابی سیستم پیشنهادی، عملکرد آن بر روی تصاویر مجموعه داده Corel مورد بررسی قرار گرفته است. همان‌طور که در شکل ۲ نشان داده شده است، بعد از خواندن تصاویر ویژگی‌های استخراج‌شده به‌صورت بردار ویژگی‌ها درمی‌آیند؛ برای مثال،

بسیاری از روش‌ها نظیر CRM [۲۵]، InfNet [۲۶]، NPDE [۲۷]، SML [۲۸]، TGLM [۲۹]، MBRM [۳۰]، JEC [۳۱]، LASSO [۳۱] و TagProp [۳۲] روی آن مورد آزمایش قرار گرفته‌اند. در این مجموعه، ۵۰۰۰ تصویر جمع‌آوری شده است که به‌طور معمول، ۴۵۰۰ نمونه از تصاویر آن برای آموزش و ۵۰۰ نمونه دیگر برای تست مدل به‌کار می‌رود. در حالت کلی، این مجموعه شامل ۵۰ گروه است که در هر گروه، ۱۰۰ تصویر متشابه مانند ساحل، هوایما و ببر وجود دارد. مجموعه دیگر مجموعه‌ای از یک فرهنگ لغات است [۲۴] که برای حاشیه‌نویسی تصاویر به‌کار گرفته می‌شود. این مجموعه شامل ۳۷۴ کلمه است که به‌طور متوسط، برای هر

پس از تعیین مقادیر داده‌های بردارهای مربوط، الگوریتم خوشه‌بندی روی داده‌ها انجام می‌شود و بردارها به‌روزرسانی خواهند شد. معیاری که برای بررسی کارایی سیستم پیشنهادی در نظر گرفته شده، تعداد تصاویر دسته‌بندی‌شده در هر خوشه است. در این قسمت، سیستم پیشنهادی از جنبه‌های مختلفی ارزیابی شده است. در ابتدا نتایج خوشه‌بندی روش پیشنهادی در مراحل مختلف بازخورد با روش K-میانگین مورد مقایسه قرار گرفته است. سپس عملکرد سیستم پیشنهادی روی مجموعه تست مورد بررسی قرار گرفته است و همچنین، کارایی سیستم پیشنهادی در مراحل مختلف بازخورد مرتبط مورد بررسی قرار گرفته است. در پایان، دقت سیستم پیشنهادی با استفاده از معیار NMI مورد ارزیابی قرار گرفته است. نخست نتایج به‌دست‌آمده از روش پیشنهادی با نتایج به‌دست‌آمده از خوشه‌بندی K-میانگین مورد مقایسه قرار گرفته است. شکل ۳ نتایج خوشه‌بندی سیستم پیشنهادی بعد از اولین بازخورد نشان می‌دهد و همان‌طور که مشاهده می‌شود، داده‌ها با توزیع مناسبی در خوشه‌ها قرار نگرفته‌اند.



شکل (۳): نتایج خوشه‌بندی با استفاده از سیستم پیشنهادی پس از اولین بازخورد

شکل ۴ نتایج خوشه‌بندی را بعد از دومین بازخورد نشان می‌دهد که نسبت به شکل ۳ توزیع داده‌ها در خوشه‌ها مناسب‌تر از حالت قبل است.

بردار ویژگی استخراج‌شده برای متوسط رنگ حاوی ۴۹ ویژگی است. جدول ۱ ویژگی‌های رنگ استخراج‌شده مربوط به پنج تصویر اول پایگاه داده را نشان می‌دهد. بردار دیگری که مورد استفاده قرار می‌گیرد، بردار داده‌های برچسب‌خورده توسط کاربر است. این بردار شامل پنج ستون است که برچسب‌های مربوط به هر تصویر را نشان می‌دهند. به دلیل اینکه هر تصویر ممکن است محتویات متفاوتی داشته باشد، برای هر تصویر پنج برچسب در نظر گرفته شده است. جدول ۲ نمونه‌ای از تصاویر برچسب‌گذاری‌شده توسط کاربر را نشان می‌دهد. شماره هر برچسب با شماره ردیف کلمه توصیف‌کننده تصویر در مجموعه فرهنگ لغات مرتبط است.

## ۲.۶. آزمایش‌ها

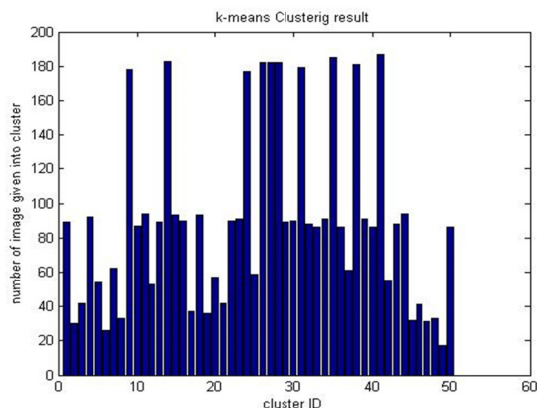
تصاویر می‌توانند دارای هیچ یا حداکثر پنج برچسب باشند که در ابتدا تعدادی از تصاویر به‌طور تصادفی، توسط کاربر برچسب می‌خورند و دیگر تصاویر با استفاده از الگوریتم پیشنهادی به‌طور متوسط، ۳.۵ برچسب اختیار می‌کنند. همان‌طور که قبلاً توضیح داده شد، از دو ماتریس M و C برای خوشه‌بندی تصاویر استفاده می‌شود. این دو ماتریس ابتدا توسط کاربر مقداردهی اولیه می‌شوند و بر طبق الگوریتم در طی اجرا به‌روزرسانی می‌شوند. جداول ۳ و ۴ تعدادی از مقادیر اولیه ماتریس‌های M و C را نشان می‌دهند.

جدول (۳): مقادیر اولیه ماتریس M

شماره تصویر اول	شماره تصویر دوم
۴	۴۲
۱۵	۸
۱۰۲	۱۷۴

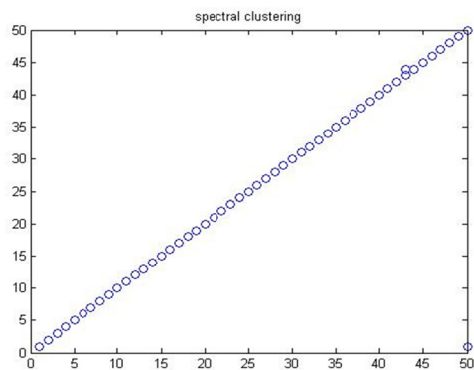
جدول (۴): مقادیر اولیه ماتریس C

شماره تصویر اول	شماره تصویر دوم
۴	۴۶۶۰
۱۵	۱۵۵
۱۰۲	۲۲۷۳

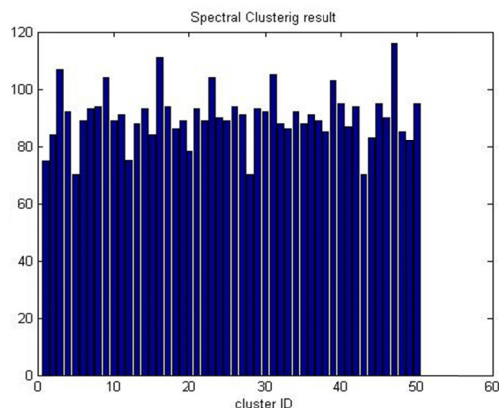


شکل (۶): نتایج خوشه‌بندی با استفاده از الگوریتم k-میانگین

در این بخش، برای بررسی دقت روش پیشنهادی عملکرد آن بر روی مجموعه داده تست مورد بررسی قرار گرفته است. همان‌طور که در بخش ۱.۶ گفته شد تعداد داده‌های مجموعه تست ۵۰۰ عدد می‌باشد. شکل ۷ نتایج اعمال روش پیشنهادی روی مجموعه داده تست را نشان می‌دهد و شکل ۸ نشان‌دهنده اعمال روش K-میانگین بر روی مجموعه داده تست است. در این نمودارها محور افقی نشان‌دهنده خوشه مربوط به داده مورد بررسی و محور عمودی خوشه تشخیص داده شده توسط روش مورد بررسی است. نتایج مربوط به نمودارهای موجود در شکل‌های ۷ و ۸ نشان‌دهنده دقت بالای سیستم پیشنهادی نسبت به الگوریتم K-میانگین در تشخیص خوشه مربوط به داده مورد بررسی است.

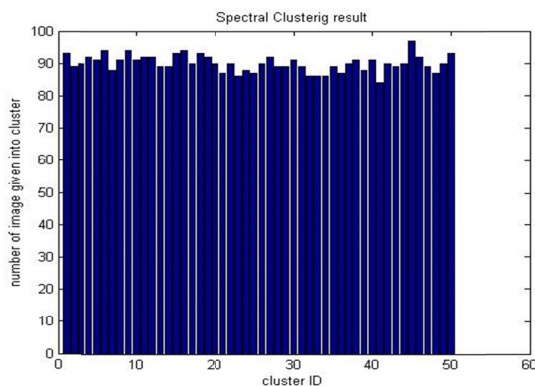


شکل (۷): نتایج اعمال روش پیشنهادی بر روی مجموعه داده تست



شکل (۴): نتایج خوشه‌بندی با استفاده از سیستم پیشنهادی پس از دومین بازخورد

شکل ۵ نتایج حاصل از سیستم پیشنهادی را پس از بازخورد سوم نشان می‌دهد و شکل ۶ نشان‌دهنده نتایج به‌دست‌آمده از روش خوشه‌بندی K-میانگین است. همان‌طور که قبلاً گفته شد، تعداد داده‌ها در هر خوشه برابر با ۱۰۰ می‌باشد. نتایج شکل ۵ توزیع صحیح و مناسب‌تر داده‌ها را در خوشه‌های مختلف و با درصد کمی خطا نشان می‌دهد، در صورتی که نتایج شکل ۶ نشان می‌دهد که تعداد تصاویر قرار گرفته در بعضی خوشه‌ها از ماکزیمم مقدار تصاویر متعلق به آن خوشه بیشتر است و خطای بیشتری را در خوشه‌بندی نشان می‌دهد.



شکل (۵): نتایج خوشه‌بندی با استفاده از سیستم پیشنهادی پس از سومین بازخورد

نشان‌دهنده دقت بالاتر روش مورد بررسی است. جدول ۶ نتایج حاصل از اعمال NMI را بر روی سیستم پیشنهادی و الگوریتم K-میانگین نشان می‌دهد.

جدول (۶): نتایج حاصل از اعمال معیار NMI بر روی سیستم پیشنهادی

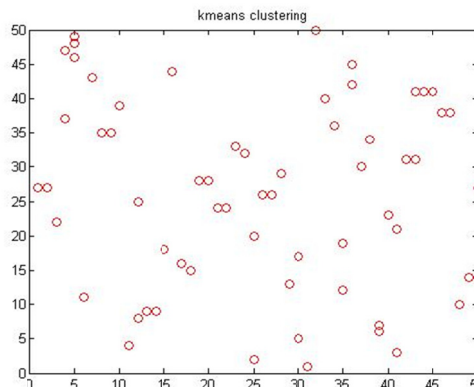
در مقایسه با الگوریتم K-میانگین

NMI	
۰.۹۹۰۶	سیستم پیشنهادی
۰.۹۳۷۲	K-میانگین

همان‌طور که از نتایج جدول ۶ مشخص است، دقت روش پیشنهادی در تشخیص برچسب‌های درست هر داده بسیار بالاست.

#### ۷. نتیجه‌گیری و کارهای آینده

با توجه به نتایج ارائه شده در بخش قبل، می‌توان گفت سیستم پیشنهادی از دقت فراوانی برخوردار است و روشی کارا در حاشیه‌نویسی تصاویر می‌باشد. از مزایای دیگر سیستم پیشنهادی می‌توان به کاهش بعد بردارهای ویژگی به دلیل استفاده از الگوریتم خوشه‌بندی طیفی اشاره کرد که پیچیدگی زمانی مناسب را در پی خواهد داشت؛ در نتیجه، می‌توان از آن برای پایگاه داده‌های بزرگ هم استفاده کرد. همچنین، استفاده از روش باز خورد مرتبط می‌تواند باعث بهبود کارایی سیستم شود. از جمله کارهایی که برای تحقیقات آینده پیشنهاد داده می‌شود، استفاده از ماتریس شباهت فازی در سیستم پیشنهاد داده شده است.



شکل (۸): نتایج اعمال الگوریتم k-میانگین بر روی مجموعه داده تست

همان‌طور که گفته شد، در روش پیشنهادی از ۳ مرحله بازخورد مرتبط استفاده می‌شود. جدول ۵ نتایج حاصل از اعمال ۳ مرحله بازخورد مرتبط را روی دقت خوشه‌بندی نشان می‌دهد. در واقع، این بخش به بررسی تأثیر بازخورد روی سیستم پیشنهادی و همچنین، الگوریتم K-میانگین می‌پردازد.

جدول (۵): تأثیر بازخورد مرتبط ترکیبی طی سه بازخورد در دقت خوشه‌بندی

بازخورد سوم	بازخورد دوم	بازخورد اول	بدون بازخورد	
٪۹۰	٪۸۸.۳	٪۸۵.۴	٪۸۰	سیستم پیشنهادی
٪۸۵.۱	٪۸۰.۳	٪۷۴.۶	٪۷۱.۵	K-میانگین

همچنین، برای بررسی بیشتر دقت سیستم پیشنهادی از معیار NMI<sup>۱</sup> استفاده شده است. این معیار یکی از پرکاربردترین معیارها برای بررسی کیفیت خوشه‌بندی است و به صورت اطلاعات مشترک بین خوشه‌های تشخیص داده شده و برچسب‌های واقعی مجموعه داده بیان می‌شود و عددی بین صفر و یک می‌باشد؛ هرچه این مقدار به ۱ نزدیک‌تر باشد،

## مراجع

- [1] Shah, B., Benton, R., Wu, Z., Raghavan, V., "Automatic and Semi-Automatic Annotation Techniques for Image", Book Chapter, In Yu-Jin Zhang, editor, Semantic-Based Visual Information Retrieval, Idea Group Publishing, 2006.
- [2] Tsai, Ch., Hung, C., "Automatically Annotating Images with Keywords: A Review of Image Annotation Systems", International Journal of Recent Patents on Computer Science, Vol.1, No. 1, pp.55-68, 2008.
- [3] Yang, H., Lee, Ch., "Image Semantics Discovery from Web Pages for Semantic-Based Image Retrieval Using Self-Organizing Maps", International Journal of Expert Systems with Applications, Vol. 34, No. 1, pp. 266-279, 2008.
- [4] Jin, W., Shi, R., Chua T. S., "A Semi-Naïve Bayesian Method Incorporating Clustering with Pair-Wise Constraints for Auto Image Annotation", The 12<sup>th</sup> ACM International Conference on Multimedia, New York, pp.336-339, 2004.
- [5] Srikanth M., Varner J., Bowden M., Moldovan D., "Exploiting Ontologies for Automatic Image Annotation", The 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Video and Image, Salvador, Brazil, pp.552-558, 2005.
- [6] Tang J., "Automatic Image Annotation and Object Detection", Ph.D. Thesis, School of Electronics and Computer Science, University of Southampton, United Kingdom, 2008.
- [7] Qiu B., Xu C., Tian Q., "Efficient Relevance Feedback using Semi-supervised Kernel-specified K-means Clustering", The 18<sup>th</sup> International Conference on Pattern Recognition, Hong Kong, Vol.3, pp.316-319, 2006.
- [8] Chiang C. C., Hung M. W., Hung Y.P., Leow W.K., "Image Annotation with Relevance Feedback Using a Semi-Supervised and Hierarchical Approach", The 3<sup>rd</sup> International Conference on Computer Vision Theory and Applications, Madeira, Portugal, Vol. 2, pp.173-175, 2008.
- [9] Liu D., Hua A., Vu K., Yu N., "Fast Query Point Movement Technique in Large Content Based Image Retrieval System", International Journal of IEEE Transaction on Knowledge and Data Engineering, Vol. 21, No. 5, pp.729-743, 2009.
- [10] Murthy V., Vamsidhar E., Kumar J., Sankara P., "Content Based Image Retrieval Using Hierarchical and K-Means Clustering Techniques", International Journal of Engineering Science and Technology, Vol. 2, No. 3, pp.209-212, 2010.
- [11] Kekre H.B, Mishra D., "Content Based Image Retrieval using Density Distribution and Mean of Binary Patterns of Walsh Transformed Color Images", International Journal on Computer Science and Engineering, Vol. 3, No.2, pp. 642-651, 2011.
- [12] Aujol J., Gilboa G., Chan T., Osher S., "Structure-Texture Image Decomposition-Modeling, Algorithms, and Parameter Selection", International Journal of Computer Vision, Vol. 67, No. 1, pp. 111-136, 2006.
- [13] Kekre H.B., Thepade S.D., Banura V.K., "Augmentation of Color Averaging Based Image Retrieval Techniques Using Even Part of Images and Amalgamation of Feature Vectors", International Journal of Engineering Science and Technology, Vol. 2, No.10, pp. 5238-5246, 2010
- [14] Ojala T., Pietikainen M., Maenpaa T., "Multi Resolution Gray Scale and Rotation Invariant Texture Analysis with Local Binary Pattern", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 7, pp. 971-987, 2002.
- [15] Yen C.Y., Cios K.J., "Image Recognition System Based on Novel Measures of Image

- Similarity and Cluster Validity*", The Journal of Neurocomputing, Vol. 72, No 1-3, pp. 401-412, 2008.
- [16] Lowe D. G., "Object Recognition from Local Scale-Invariant Features", The 7<sup>th</sup> IEEE International Conference on Computer Vision, Kerkyra ,Vol. 2, pp. 1150-1157, 1999.
- [17] Lowe D. G., "Distinctive Image Features from Scale-Invariant Key points", International Journal of Computer Vision, Vol. 60, No. 2, pp. 91-110, 2004.
- [18] Nissim, K., Harel, E., "A Texture Based Approach to Defect Analysis of Grapefruits", The 19<sup>th</sup> IEEE International Parallel and Distributed Processing Symposium, Vol. 1, Switzerland, 1997.
- [19] Liu G.H, Yang J.Y, "Image Retrieval Based on the Texton Co-Occurrence Matrix", Journal of Pattern Recognition, Vol. 41, No. 12, pp. 3521-3527, 2008.
- [20] Jones J., Palmer L., "An Evaluation of the Two Dimensional Gabor Filter Model of Simple Receptive Fields in Cat Striate Cortex", Journal of Neurophysiology, Vol. 58, No. 6, pp. 1233-1258, 1987.
- [21] Daugman J., "Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and Compression", IEEE Transactions on Speech Signal Processing, Acoustics. Vol. 36, No. 7, pp.1169-1179, 1988.
- [22] Klein D., Kamvar S. D., Manning C. D., "From Instance-level Constraints to Space-level Constraints: Making the Most of Prior Knowledge in Data Clustering", The 19<sup>th</sup> International Conference on Machine Learning, Sydney, Australia, pp. 307-314, 2002.
- [23] Chen W., Feng G., "Spectral Clustering: A Semi-Supervised Approach", Journal of Neurocomputing, Vol. 77, No. 1, pp. 229-242, 2012.
- [24] Duygulu P., Barnard K., DeFreitas J.F.G., Forsyth D.A., "Object Recognition as Machine Translation: Learning a Lexicon for a Fixed Image Vocabulary", The 7<sup>th</sup> European Conference on Computer Vision, pp. 97-112, Copenhagen, 2002.
- [25] Lavrenko V., Manmatha R., Jeon J., "A Model for Learning the Semantics of Pictures", The 16<sup>th</sup> Conference on Advances in Neural Information Processing Systems, Vol. 1, pp. 2, 2003.
- [26] Metzler D., Manmatha R., "An Inference Network Approach to Image", The 3<sup>rd</sup> International Conference on Image and Video Retrieval, Dublin, Ireland, Vol. 3115, pp. 42-50, 2004.
- [27] Yavlinsky A., Schofield E., Ruger S., "Automated Image Annotation Using Global Features and Robust Nonparametric Density Estimation", The 4<sup>th</sup> International Conference on Image and Video Retrieval, Singapore, Vol. 3568, pp. 507-517, 2005.
- [28] Carneiro G., Chan A. B., Moreno P. J., Vasconcelos N., "Supervised Learning of Semantic Classes for Image Annotation and Retrieval, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 29, No. 3, pp. 394-410, 2007.
- [29] Liu J., Li M., Liu Q., Lu H., Ma S., "Image Annotation via Graph Learning", Journal of Pattern Recognition, Vol. 42, No. 2, pp. 218-228, 2009.
- [30] Feng S. L., Manmatha R., Lavrenko V., "Multiple Bernoulli Relevance Models for Image and Video Annotation", The International Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, Vol. 2, pp. 1002-1009, 2004.
- [31] Makadia A., Pavlovic V., Kumar S., "A New Baseline for Image Annotation", The 10<sup>th</sup> European Conference on Computer Vision, Marseille, France, pp. 316-329, 2008.
- [32] Guillaumin M., Mensink T., Verbeek J., Schmid C., "Tagprop: Discriminative Metric Learning in Nearest Neighbor Models for Image Auto-Annotation", The 12<sup>th</sup> International IEEE



Conference on Computer Vision, Kyoto, pp.  
309-316, 2009.