

Vol. 12/ No. 45/Autumn 2022

Research Article

# Extraction of Multiple Hybrid Features to Reduce the Semantic Vacuum with the Semi-Supervised Classification

Mahdi Jalali, Assistant Professor <sup>1,3</sup>  | Tohid Sedghi, Assistant Professor <sup>2,3</sup> 

<sup>1</sup>Department of Electrical Engineering, Naghadeh Branch, Islamic Azad University, Naghadeh, Iran, [Mahdi.jalali@iauna.ac.ir](mailto:Mahdi.jalali@iauna.ac.ir)

<sup>2</sup>Department of Electrical Engineering, Urmia Branch, Islamic Azad University, Urmia, Iran, [t.sedghi@iaurmia.ac.ir](mailto:t.sedghi@iaurmia.ac.ir)

<sup>3</sup>Microwave and antenna research center, Urmia Branch, Islamic Azad University, Urmia, Iran

**Correspondence**

Tohid Sedghi, Assistant Professor of Electrical Engineering, Urmia Branch, Islamic Azad University, Urmia, Iran, [t.sedghi@iaurmia.ac.ir@iau.ac.ir](mailto:t.sedghi@iaurmia.ac.ir@iau.ac.ir)

**Received:** 28 February 2022

**Revised:** 11 April 2022

**Accepted:** 12 May 2022

## Abstract

In this paper, for the classification of images, the observed cooperative classification method is proposed with the aim of reducing the semantic vacuum. Because most classification methods are sensitive to cluster centers at initialization, the algorithm converges optimally locally if the quantification is not done correctly. It is also very difficult to combine the results of the classification due to the fact that the labels of the work centers are not clear. Supervised semi-classification is used to solve these problems. To achieve the highest performance, the system classification results are combined with different color space and similarity criteria with multiple features as a semi-supervised cooperative. When the number of features is effective, the relevant feedback is used for its semi-regulatory classification. One of the most important parts of an image retrieval system and a classification algorithm is to determine the appropriate similarity between images. In this study, two methods of classification of cases have been used, which include k-NN and PNN classification, which according to the results in all proposed methods, a better response from k-NN classification than PNN has been observed. The proposed algorithm is very suitable for classifying large image databases due to the reduction of time complexity. Also, in image retrieval in different color spaces and using different similarity criteria, different classifications are obtained. Better results can be achieved if the classification results are combined. Recognition rate on the image data used in this study, on which the pyramidal histogram algorithm of oriented gradients is applied, has the highest rate compared to other proposed methods and is 98.52%. Experiments on the Corel image database show that the combined method increases the classification accuracy by an average of about 20% in the combined method.

**Keywords:** Semi-supervised classification, image retrieval, feature, annotation

## Highlights

- Extraction of multiple hybrid features to reduce semantic vacuum using related feedback for Semi-Supervised classification
- Sampling points from the greatest distance from the center of gravity and keeping each of the image vector modes constant

**Citation:** M. Jalali and T. Sedghi, "Extraction of Multiple Hybrid Features to Reduce the Semantic Vacuum with the Semi-Supervised Classification," *Journal of Southern Communication Engineering*, vol. 12, no. 45, pp. 31–44, 2022, [doi:10.30495/jce.2022.691134](https://doi.org/10.30495/jce.2022.691134), (in Persian).

## مقاله پژوهشی

## استخراج ویژگی های چندگانه ترکیبی برای کاهش خلاً معنایی با طبقه‌بندی نیمه نظارتی

مهدی جلالی<sup>۱</sup> | توحید صدقی<sup>۲\*</sup> ID

## چکیده:

در این مقاله برای طبقه‌بندی تصاویر، روش طبقه‌بندی تعاونی نظارت‌شده با هدف کاهش خلاً معنایی پیشنهاد می‌شود. اکثر روش‌های طبقه‌بندی به مقاردهی اولیه به مراکز خوشه حساس هستند و اگر به درستی مقاردهی انجام نشود الگوریتم به بهینه محلی همگرا می‌شود. همچنین ترکیب نتایج طبقه‌بندی به دلیل مشخص نبودن برچسب مراکز کار بسیار مشکلی است. برای برطرف کردن این مشکلات از طبقه‌بندی نیمه نظارت شده استفاده می‌شود. برای دستیابی به بالاترین کارایی، نتایج طبقه‌بندی سیستم با فضای رنگ و معیار شباهت متفاوت با ویژگی‌های متعدد به صورت تعاونی نیمه‌نظارتی با هم ترکیب می‌شوند. در شرایطی که تعداد ویژگی‌ها مؤثر هستند، از بازخورد مرتبط برای طبقه‌بندی نیمه نظارتی استفاده می‌شود. در این پژوهش از دو روش طبقه‌بندی حالات استفاده شده است که شامل طبقه‌بندی k-NN و PNN است که با توجه به نتایج در همه روش‌های پیشنهاد شده، از طبقه‌بندی k-NN پاسخ بهتری نسبت به PNN مشاهده شده است. الگوریتم پیشنهادی به دلیل کاهش پیچیدگی زمان، برای طبقه‌بندی پایگاه داده‌های بزرگ تصاویر بسیار مناسب است. نرخ بازشناسی بر داده‌های تصویری استفاده شده در این تحقیق که الگوریتم هیستوگرام هرمی گرادین‌های جهت‌دار بر آن‌ها اعمال شده، دارای بالاترین نرخ نسبت به دیگر روش‌های پیشنهادی بوده و ۹۸٫۵۲٪ است. آزمایش‌ها روی پایگاه داده تصاویر Corel نشان می‌دهند که روش ترکیبی افزایش دقت طبقه‌بندی به‌طور میانگین در روش ترکیبی حدود ۲۰٪ است.

کلیدواژه‌ها: طبقه‌بندی نیمه نظارتی، بازیابی تصاویر، ویژگی، حاشیه-نویسی.

<sup>۱</sup>استادیار گروه برق، دانشکده فنی مهندسی، واحد نقده، دانشگاه آزاد اسلامی، نقده، ایران، Mahdi.jalali@iauna.ac.ir

<sup>۲</sup>استادیار گروه برق، واحد ارومیه، دانشگاه آزاد اسلامی، ارومیه، ایران Tohid.Sedghi@iau.ac.ir & Sedghi.tohid@gmail.com

<sup>۳</sup>مرکز تحقیقات میکروویو و آنتن، واحد ارومیه، دانشگاه آزاد اسلامی، ارومیه، ایران

نویسنده مسئول

\*توحید صدقی، استادیار، گروه مهندسی برق، واحد ارومیه،

دانشگاه آزاد اسلامی، ارومیه، ایران

مرکز تحقیقات میکروویو و آنتن، واحد ارومیه، دانشگاه آزاد

اسلامی، ارومیه، ایران

Tohid.Sedghi@iau.ac.ir &amp; Sedghi.tohid@gmail.com

تاریخ دریافت: ۹ اسفند ۱۴۰۰

تاریخ بازنگری: ۲۲ فروردین ۱۴۰۱

تاریخ پذیرش: ۲۲ اردیبهشت ۱۴۰۱

<https://doi.org/10.30495/jce.2022.691134>

## ۱-مقدمه

امروزه با افزایش روزافزون تصاویر دیجیتال نیاز به روش‌های طبقه‌بندی و جستجوی مؤثر برای بایگانی بزرگ تصاویر ضروری است [۱]. تنوع مجموعه تصاویر موجود و انواعی از کاربردها (تبلیغات، امنیت [۲] پزشکی [۳]، تصاویر ماهواره‌ای [۴]، صنعتی، شخصی و تجاری [۵-۹]) باعث توسعه وسیعی از روش‌های فوق شده است [۱۰-۱۴]. بدین منظور سیستم طبقه‌بندی برای پاسخ

به این نیازها معرفی شد. به کمک ویژگی‌های بصری مثل رنگ و بافت و شکل تصاویر به گروه‌هایی با مفاهیم مشابه دسته‌بندی می‌شوند [۱۵-۱۷]. در حالت کلی پیوند مستقیمی بین مفاهیم سطح بالا و ویژگی‌های سطح پایین وجود ندارد. ویژگی‌های سطح پایین اغلب در توصیف مفاهیم معنایی سطح بالا در ذهن کاربران ناتوان هستند. این تفاوت فاصله بین قدرت توصیفگرهای سطح پایین و معانی غنی کاربر به‌عنوان فاصله معنایی تعریف می‌شود [۱۰]. برای کاهش فاصله معنایی، بازخورد مرتبط<sup>۱</sup> (RF) [۸] استفاده می‌شود. بازخورد مرتبط یک فرایند خودکار است که سعی در شناخت خواسته‌های کاربر در حین اجرا دارد. در روش‌های بازیابی متن و سیستم بازیابی تصویر بر مبنای محتوا با قصد آوردن کاربر در حلقه بازیابی برای کاهش فاصله معنایی بین اینکه پرس‌وجو چه می‌گوید (ویژگی‌های سطح پایین) و کاربر چه فکر می‌کند، بازخورد مرتبط معرفی شد. با تعامل داشتن با کاربر، بازخورد مرتبط باعث بهبود کارایی سیستم طبقه‌بندی به‌طور قابل توجهی می‌شود [۶].

در این مقاله، برای طبقه‌بندی تصاویر از الگوریتم طبقه‌بندی k-means نیمه نظارت‌شده به همراه PNN<sup>۲</sup> استفاده می‌شود [۷]. به دلیل استفاده از این روش هم‌زمان در بازیابی، کاهش فاصله معنایی را خواهیم داشت و افزایش دقت طبقه‌بندی به‌طور میانگین در روش ترکیبی حدود ۲۰٪ است در صورتی هر سه روش به‌طور جداگانه افزایش دقتی حدود ۹-۱۱٪ را دارند. برای بهبود دقت طبقه‌بندی ترکیبی از سه مرحله بازخورد استفاده می‌شود. این سه مرحله عبارت‌اند از:

- از تصاویری که کاربر نامربوط به خوشه می‌داند، با استفاده از بازخورد مرتبط طولانی مدت به خوشه درست هدایت می‌شوند.
- از تصاویری که کاربر مربوط به خوشه می‌داند، با استفاده از بازخورد مرتبط و تغییر وزن ویژگی و وزن ویژگی خوشه‌ها تخمین زده شده و یک معیار شباهت چندگانه ارائه می‌شود.
- با استفاده از بازخورد مرتبط مراکز خوشه‌ها به‌صورت دقیق تعیین می‌شوند. آزمایش‌ها روی ۱۰۰۰ تصویر پایگاه داده Corel نشان می‌دهد که کارایی روش ترکیبی پیشنهادی از تک‌تک روش‌ها بهتر است و منجر به کاهش فاصله معنایی می‌گردد.

## ۲-مروری بر کارهای انجام شده

اکثر محققان برای گسترش روش‌های بازخورد مرتبط برای رسیدن به کارایی بالاتر کوشش کرده‌اند. چگونگی استفاده از کاربرد نظرات کاربران در قالب روش‌های یادگیری یکی از مهم‌ترین مراحل سیستم طبقه‌بند است. سیستم طبقه‌بند فاصله بین تصویر پرس‌وجو و تصاویر پایگاه داده را بر اساس معیار شباهت محاسبه می‌کند و t تصویر را که کمترین فاصله با تصویر پرس‌وجو را دارد برای بازخورد به کاربر برمی‌گرداند. کاربر برخی از این تصاویر را مربوط و برخی دیگر را نامربوط تشخیص می‌دهد. برای یادگیری ذهنیت نظرات کاربر یک الگوریتم یادگیری نیاز است. در حالت کلی روش‌های یادگیری بازخورد مرتبط را می‌توان به دو دسته کلی روش‌های باناظر و روش‌های نیمه نظارت‌شده یا بدون ناظر تقسیم کرد. بازخورد مرتبط باناظر رابطه نزدیکی با مسائل شناسایی الگو روش‌های یادگیری ماشین دارد. از آنجا که این سیستم‌ها توسط بازخورد کاربر کنترل می‌شوند سیستم HCl<sup>۳</sup> نامیده می‌شوند که سعی در شبیه‌سازی مدل درک انسان با یک روش تحلیل جداکننده غیرخطی دارند [۱۲]. سه روش پایه برای یادگیری نظر کاربر، تغییر نقطه پرس‌وجو<sup>۴</sup> (QPM) [۱۲] و تغییر وزن ویژگی<sup>۵</sup> [۱۵][۱۶] و استنتاج بیزین<sup>۶</sup> (BI) [۱۹] [۱۶] هستند. چون محققان همواره به دنبال بالاترین استفاده از بازخوردهای کاربر هستند پس می‌توان این روش‌ها را به‌صورت ترکیبی استفاده کرد.

محققان سعی در ترکیب مدل‌های بازخورد مرتبط مختلف برای به دست آوردن روش بهینه دارند. برای نمونه در [۱۶] ترکیبی از سه روش Re-weighting, QPM و BI را با استفاده از یادگیری تقویتی برای به دست آوردن بهترین ترکیب این سه روش در پاسخ به هر پرس‌وجو به‌کاربرده است. در [۱۴] BALAS ترکیبی از QPM و BI را با تنظیم درجه مربوط یا نامربوط بودن تصاویر به‌صورت آنلاین توسط تخمین تابع چگالی احتمال استفاده کرده‌است. سیستم‌های نیمه نظارت شده و بدون ناظر به دلیل کاهش

<sup>1</sup> Relevance Feedback (RF)

<sup>2</sup> Probabilistic Neural Network

<sup>3</sup> Human-Computer Interaction

<sup>4</sup> Query Vector Modification

<sup>5</sup> Feature Relevance Estimation or Re-Weighting

<sup>6</sup> Bayesian Inference(BI)

کارهای کاربر امروزه بیشتر مورد توجه قرار گرفته‌اند. سیستم‌های<sup>۱</sup> MCI یک روش مؤثر برای بازیابی خودکار است [۲۱]. بنابراین همواره استفاده از نظرات چندین کاربر کارایی بسیار بالاتری دارد. بنابراین از نظر نگهداری اطلاعات بازخوردها، روش‌های بازخورد مرتبط را می‌توان به دو دسته یادگیری کوتاه‌مدت و یادگیری بلندمدت تقسیم کرد.

در سال‌های اخیر توجه به روش‌های طبقه‌بندی در حال افزایش است. ولی با توجه به نقش مؤثر بازخورد مرتبط در سیستم‌های طبقه‌بندی تعداد کمی از مقالات به روش‌های بازخورد مرتبط برای طبقه‌بندی پرداخته‌اند. CLUE [۲۱] یکی از مقالاتی که ابتدا از تئوری گراف و روش Ncut برای طبقه‌بندی 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 نداشت مستقیمی بین ویژگی‌های سطح پایین و مفاهیم سطح بالا ایجاد کند. همچنین زمان محاسباتی زیادی نیاز دارد. در این پژوهش در حالت کلی مطالعات انجام شده به دو بخش تقسیم می‌گردد. یک بخش مبتنی بر کلاس‌بندی حالات است که اطلاعات سطح بالا در مورد اشیاء متحرک را فراهم می‌کند. بخش دیگر مبتنی بر تخمین موقعیت است که حرکت سه‌بعدی واقعی تصویر را در نظر می‌گیرد. طبقه‌بندی حالت، مبتنی بر ظاهر تصویر بوده و روش‌های بسیاری برای استخراج ویژگی را شامل می‌گردد، مثل حرکت، تحلیل فوری و اطلاعات لبه تصویر. سپس حالاتی که با تخمین موقعیت همراه باشند، مورد بررسی قرار گرفته‌اند. در دسته‌بندی دیگری، ویژگی‌ها به سه گروه شامل ویژگی‌های سطح بالا، ویژگی‌های سطح پایین و ویژگی‌های سه‌بعدی تقسیم شده‌اند.

الگوریتم‌هایی که ویژگی سطح بالا را به کار می‌گیرند، اغلب با تکیه بر نشان‌گرهای رنگی برای استخراج ویژگی و مکان اتصالات و مکان‌های مهم روی تصویر است. مرکز ثقل تصویر به‌عنوان یک ویژگی سطح بالا مطرح می‌گردد. در برخی از مطالعات، ویژگی‌هایی مانند برآمدگی یا فرورفتگی‌های تصویر برای آشکارسازی استفاده می‌گردد. اتصالات میان تصاویر ویژگی‌های به هم‌فشرده‌ای هستند که محاسبات پرسرعتی را برای سنجش میزان شباهت فراهم می‌نمایند. اما ویژگی‌های سطح بالا در مقابل مانع ضعیف هستند و ضمناً هیچ‌کدام از ویژگی‌های بالا در پس‌زمینه‌های بی‌نظم و شلوغ درست عمل نمی‌کنند. در کل استخراج ویژگی سطح بالا معمولاً دشوار است، لذا ویژگی‌های سطح پایین و ویژگی‌های سه‌بعدی مورد استفاده قرار می‌گیرند [۲۳]. منحنی پیرامونی لبه ویژگی‌های کلی هستند که در هر روش مبتنی بر مدل مورد استفاده قرار می‌گیرند. روش‌های مبتنی بر لبه در پس‌زمینه‌های به‌هم‌ریخته مؤثر نخواهند بود. در پژوهش [۲۴] برای استخراج ویژگی از تصویر، روشی تحت عنوان فضای مقیاس منحنی<sup>۲</sup> (CSS) مطرح شده است. روش CSS از سه مرحله شامل بخش‌بندی تصویر<sup>۳</sup>، استخراج ویژگی و بازشناسی تشکیل می‌یابد. مراحل کار به‌صورت زیر هستند:

- منحنی پیرامونی تصویر به‌صورت باینری (دودویی) مشخص می‌گردد.
  - فیلتر پایین‌گذر با هسته‌ای که به آن مقادیر (۰٫۲۵، ۰٫۵، ۰٫۲۵) داده شده است، اعمال می‌گردد.
- تعداد دفعات لازم برای حذف هر فرورفتگی بین دو قله را مشخص می‌کنند. از تصویر CSS به تصویر آمده، مختصات فرورفتگی بین قله‌ها به‌صورت رابطه (۱) تعیین می‌شود:

$$\{(u_i, \sigma_i)_{i=1, \dots, N}^{original}\} = \{(u_1, \sigma_1), \dots, (u_N, \sigma_N)\} \quad (1)$$

که در آن  $N$  تعداد فرورفتگی‌های آشکار شده در تصویر CSS،  $1 \leq i \leq N$  است.  $\sigma_i$  نشان‌دهنده تعداد دفعات لازم برای حذف فرورفتگی  $i$ ام و  $u_i$  بیان‌گر مقدار منحنی پیرامونی در فرورفتگی  $i$ ام است. این ویژگی‌ها به شیوه دیگری نیز بیان می‌گردند. به این ترتیب که مقادیر بزرگ‌ترین قله، استاندارد شده و رابطه (۱) به‌صورت رابطه (۲) بازنویسی می‌شود:

$$\left\{ \{(u_i, \sigma_i)\}_{i=1, \dots, N}^{t-Normalized} \mid t = 1, 2, \dots, P \right\} \quad (2)$$

که در آن  $P$  تعداد فرورفتگی‌های عمیق بین قله‌های منحنی پیرامونی است. اگر تعداد فرورفتگی‌ها صفر باشد، مقدار  $P$  برابر با یک قرار داده خواهد شد. ویژگی‌های نسبت به چرخش، انتقال و تغییر اندازه حساس نیست. در پژوهش [۲۶] اساس سیستم‌های بازشناسی به دو دسته تقسیم می‌شوند. سیستم‌های مبتنی بر داده<sup>۴</sup> و سیستم‌های مبتنی بر مدل<sup>۵</sup>. هدف نهایی هر دو سیستم،

<sup>1</sup> Machine - Computer Interaction

<sup>2</sup> Curvature Scale Space

<sup>3</sup> Segmentation

<sup>4</sup> Data-Driven

<sup>5</sup> Model-Driven

نگاشت تصویر به خروجی مطلوب است. یک مرحله پیش پردازش صورت می گیرد و ناحیه مورد علاقه انتخاب می گردد، سپس ویژگی های آن استخراج می گردند و این ویژگی ها در مرحله بعد تا حد امکان کاهش می یابند. سپس به عنوان ورودی به مرحله بازشناسی وارد می گردند. در سیستم های مبتنی بر مدل، ورودی سیستم، الگوهای معتبر تصویر است. از آنجایی که مدل، نگاشتی است که فضای پارامترها را به فضای شامل الگو می نگارد، یک روش مناسب برای تخمین پارامترها مورد نیاز است. ویژگی ها برای تخمین پارامترهای مدل جدید بر مبنای موارد مشاهده شده در تصویر، استخراج می شوند. این ویژگی های محلی به حالت کنونی مدل وابسته هستند. ویژگی ها، هم در سیستم های داده و هم در سیستم های مدل مورداستفاده قرار می گیرند.

در حالت مبتنی بر داده، تصویر دقیقاً با ویژگی ها نشان داده می شود و تلاش می شود که متغیرهایی که برای بازشناسی مهم نیستند، حذف گردند. به دلیل اهمیت بالای ویژگی های تصویر برای بازشناسی، این روش یا با تکیه بر اطلاعات رنگ پوست یا حذف پس زمینه از تصویر و یا با باینری کردن تصویر انجام می گیرد. در این پژوهش نیز پس از انتخاب ناحیه، ویژگی ها استخراج می گردند و سپس با بهره گیری از این ویژگی ها یا حالت نهایی، بازشناسی می گردد و یا پارامترهای مدل به درستی تخمین زده می شود.

در مرجع [۲۴] ویژگی لبه از تصویر استخراج شده است. معیار شباهت بین تصاویر در این مرجع فاصله چامفر است که از رابطه (۳) به دست می آید:

$$C(X, Y) = \frac{1}{|X|} \sum_{x \in X} \min_{y \in Y} \|x - y\| \quad (3)$$

که در آن  $X$  و  $Y$  تصاویر و  $x$  و  $y$  نیز به ترتیب نقاط مرزی تصاویر  $X$  و  $Y$  هستند. سپس بازشناسی تصویر با معیار فاصله  $HD^1$  و یک حافظه تصویری<sup>۲</sup> برای ذخیره سازی تصاویر، انجام می گیرد. طبق این مرجع پیکسل های رنگ پوست در محدوده معینی از فضای HSI قرار دارند. محدوده هایی که پس از بررسی شرایط نوری طبیعی و مصنوعی برای پیکسل ها تعیین شده است، عبارت اند از:

$$\begin{cases} 0 \leq hue \leq 30 \\ 60 \leq sat \leq 160 \end{cases} \quad \text{برای نور مصنوعی} \quad \begin{cases} 0 \leq hue \leq 15 \\ 20 \leq sat \leq 120 \end{cases} \quad \text{برای نور طبیعی}$$

یعنی اگر پیکسل ها در محدوده های مشخص شده در بالا قرار گرفته باشند، متعلق به پوست انسان هستند. سپس قسمت های هم رنگ از کل تصاویر جدا می گردند و در blob هایی تقسیم بندی می شوند. زیرا فرض می شود که در فاصله نزدیکتری نسبت به بقیه اعضا در تصویر قرار دارد. دو مجموعه  $A = \{a_1, a_2, \dots, a_m\}$  و  $B = \{b_1, b_2, \dots, b_n\}$  را در نظر بگیرید. فاصله  $HD$  بین آن ها به صورت زیر محاسبه می گردد:

$$H(A, B) = \max(h(A, B), h(B, A)) \quad (4)$$

که:

$$h(A, B) = \max_{a \in A, b \in B} (\min \|a - b\|) \quad (5)$$

$HD$  جهت دار از مجموعه  $A$  به مجموعه  $B$  نامیده می شود. در واقع  $HD$  نقاط مجموعه  $A$  را بر اساس نزدیکی شان به نقاط مجموعه  $B$  مرتب می کند. معیار فاصله منهای<sup>۳</sup> برای فاصله بین هر دو نقطه استفاده می گردد. برای کاهش اثر نویز رابطه  $h$  به صورت زیر اصلاح می گردد:

$$H_k(A, B) = k^{th} \min(\|a - b\|) \quad a \in A \quad (6)$$

که آن را  $HD$  جزئی می نامیم. لذا  $HD$  جزئی دو طرفه به صورت زیر خواهد بود:

$$H_{kl}(A, B) = \max(h_k(A, B), h_l(B, A)) \quad (7)$$

در این مرحله تطابق با استفاده از  $HD$  و حافظه تصویری انجام می شود. لبه های آن ها استخراج شده، ویژگی آن ها محاسبه شده و در حافظه ذخیره شده است. سپس فاصله  $HD$  بین هر حالت جدیدی که اتفاق می افتد و تمام موارد ذخیره شده در حافظه محاسبه می گردد و کمترین فاصله متناظر با شبیه ترین حالت در نظر گرفته می شود.

<sup>1</sup> Housdorff Distance

<sup>2</sup> Visual Memory

<sup>3</sup> Manhattan

### ۳- طبقه‌بندی و کلاس‌بندی

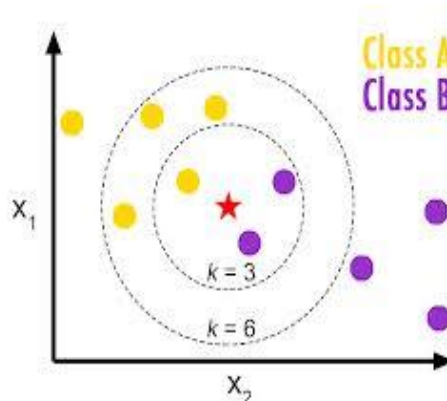
در این بخش از تحقیق، با استفاده از طبقه‌بندها به بازشناسی پرداخته می‌شود. در مرحله بازشناسی از دو نمونه کلاس‌بند شبکه‌های احتمالاتی<sup>۱</sup> (PNN) و کلاس‌بند  $k$  نزدیکترین همسایه<sup>۲</sup> (k-NN) استفاده می‌شود و عملکرد هر یک در بخش نتایج با هم مقایسه می‌شود. یک دسته الگو خانواده‌ای از الگوها هستند که دارای ویژگی‌های مشترکی هستند.

#### ۳-۱- کلاس‌بندی با استفاده از شبکه‌های احتمالاتی (PNN)

شبکه‌های شعاع مبنا<sup>۳</sup> (RBF)، یک نمونه از شبکه‌های عصبی است که دارای نورون‌های فراوان است. این شبکه‌ها زمانی که بردارهای آموزشی بسیار زیاد باشد، به علت کوتاه بودن زمان طراحی کارایی خوبی دارد. یکی از نمونه‌های شبکه‌های شعاع مبنا، شبکه‌های احتمالاتی است. شبکه‌های شعاع مبنا از دو لایه تشکیل شده‌اند. شبکه‌های عصبی احتمالاتی، احتمال عضویت یک نمونه به یک کلاس و به عبارت دیگر تابع چگالی را تخمین می‌زند. از این رو به آن PNN می‌گویند. شبکه‌های عصبی احتمالاتی، در مسائل طبقه‌بندی کاربرد فراوان دارند و شامل دو لایه است. زمانی که یک بردار ورودی به شبکه اعمال می‌شود، لایه اول، به محاسبه فاصله بردار ورودی از بردارهای ورودی آموزشی می‌پردازد و به این ترتیب برداری را فراهم می‌کند که عناصر آن تعیین‌کننده میزان فاصله بین بردار ورودی و بردار آموزشی هستند. لایه دوم با استفاده از خروجی لایه اول، برداری از احتمال‌ها را به‌عنوان خروجی شبکه تولید می‌کند. در نهایت یک تابع انتقال رقابتی موجود در لایه دوم، حداکثر مقدار احتمال‌ها را از بردار احتمال‌ها انتخاب می‌کند و به ازای آن خروجی یک و به ازای بقیه احتمال‌ها، خروجی صفر تولید می‌کند. در ساختار شبکه عصبی احتمالاتی، بردارهای ورودی، دارای  $k$  کلاس متفاوت هستند. بنابراین هر بردار هدف دارای  $k$  کلاس است که یکی از آن‌ها یک و بقیه صفر هستند. حاصل کار به‌صورت عنصر به عنصر در بایاس ضرب شده و به تابع انتقال اعمال می‌شود. در نتیجه این تابع به ازای بردار ورودی نزدیک به بردار آموزشی، مقدار یک را تولید می‌کند. مقدار  $\pi^2$  در لایه دوم از ضرب  $T$  در  $a^1$  به دست می‌آید و در نهایت  $\pi^2$  مقدار یک و به ازای بقیه مقادیر مقدار صفر را تولید می‌کند. بنابراین شبکه، بردارهای ورودی را به یکی از  $k$  کلاس موجود که دارای حداکثر احتمال است، نظیر می‌کند. در کل شبکه طوری آموزش داده می‌شود که عکس‌العمل مناسبی به داده‌های غیر از داده‌های آموزشی نشان دهد [۲۲].

#### ۳-۲- طبقه‌بندی نظارتی با روش $k$ نزدیکترین همسایه (k-NN)

اساس این کلاس‌بندی این است که در آن شیء بر اساس بیشترین رأی از همسایگانش کلاس‌بندی می‌شود. در این روش رکورد جدید، بر اساس معیار فاصله از همه رکوردهای آموزشی محاسبه می‌شود. سپس  $k$  نزدیکترین همسایه‌ها مشخص می‌گردند. در انتها دسته این  $k$  همسایه بررسی شده و دسته‌ای که بیشترین تعداد دیده شدن را در این  $k$  رکورد دارد، به‌عنوان دسته رکورد جدید در نظر گرفته می‌شود.



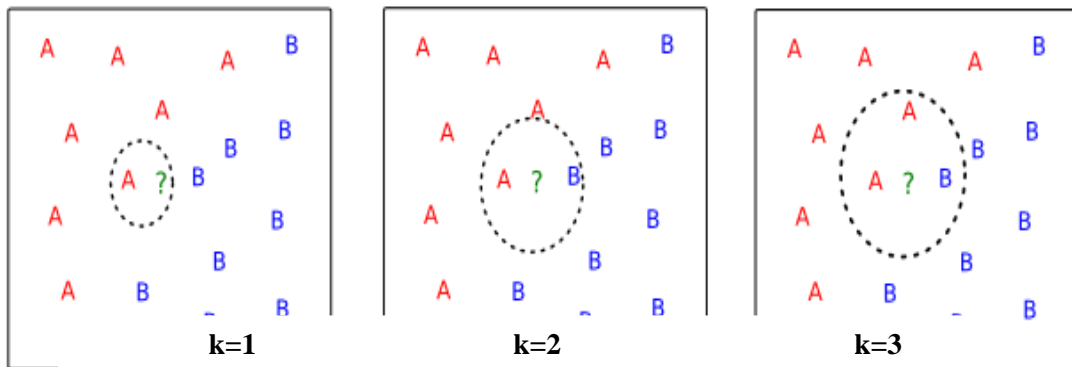
شکل ۱: الگوریتم  $k$ -NN اعمالی بر روی داده‌های آموزشی در  $k=3$  و  $k=6$

<sup>۱</sup> Probabilistic Neural Network

<sup>۲</sup> k-Nearest Neighbor

<sup>۳</sup> Radial Basis Function

در شکل ۱، هنگامی که  $k=3$  در نظر گرفته شود، دسته رکورد نامشخص که با علامت ستاره قرمز در شکل نشان داده شده است، به عنوان کلاس B که با دایره بنفش نشان داده شده است، اعلام می‌گردد. در  $k=6$ ، دسته رکورد نامشخص به عنوان کلاس A که با دایره زرد رنگ نشان داده شده است، طبقه‌بندی می‌گردد. علت این طبقه‌بندی این است که در حالت اول دو دایره بنفش و یک دایره زرد رنگ در داخل دایره داخلی قرار می‌گیرد و به همین خاطر رکورد نامشخص به عنوان کلاس B اعلام می‌شود. در حالت دوم دو دایره بنفش و چهار دایره زرد رنگ در داخل دایره بیرونی قرار می‌گیرد و چون تعداد دایره‌های زرد بیشتر است به همین خاطر رکورد نامشخص به عنوان کلاس A اعلام می‌شود. در حالتی که تنها دو دسته داشته باشیم، بهتر است مقدار  $k$  عدد فرد باشد. نکته مهم آن است که انتخاب درست  $k$  در عملکرد الگوریتم  $k$ -NN بسیار تأثیرگذار است. مثلاً در شکل ۲ برای  $k=1$ ،  $k=2$  و  $k=3$  سه جواب متفاوت گرفته می‌شود.



شکل ۲: الگوریتم  $k$ -NN که در  $k=1$  کلاس A به عنوان دسته رکورد جدید و در  $k=2$  رکورد نامشخص و در  $k=3$  کلاس A به عنوان دسته رکورد جدید

در شکل ۲، برای  $k=1$ ، کلاس A به عنوان دسته رکورد جدید خواهد بود. برای  $k=2$ ، دسته رکورد جدید نامشخص خواهد بود و برای  $k=3$ ، دسته از کلاس A خواهد بود. اگر  $k$  خیلی کوچک باشد الگوریتم به نویز حساس می‌شود و اگر  $k$  خیلی بزرگ باشد ممکن است دسته‌بندی با اشتباه صورت گیرد. معیارهای فاصله‌ای که در این الگوریتم از آن استفاده می‌شود، می‌تواند یکی از فاصله‌های اقلیدسی، city-block، کسینوسی یا همبستگی باشد. لازم به ذکر است که در مواقعی که گره ایجاد می‌شود یعنی بیشتر نقاط در یک کلاس با هم برابر هستند، از یکی از دو قانون نزدیکترین و قانون تصادفی، برای شکستن گره استفاده می‌شود. قانون نزدیکترین، یعنی از بین گره‌های به هم خورده، مجدداً کلاس نزدیکترین همسایه به عنوان کلاس خروجی انتخاب می‌شود. قانون تصادفی، یعنی برای شکستن گره، به صورت تصادفی به یکی از همسایه‌هایی که هم گره خورده‌اند، انتخاب شده و کلاس آن همسایه به عنوان کلاس خروجی بازگردانده می‌شود. مشاهده می‌شود که استفاده هم‌زمان از دو طبقه‌بند نیمه‌نظارتی PNN و  $K$ -NN تأثیر مطلوبی در کاهش خلأ معنایی دارد که در این تحقیق از این دو طبقه‌بند استفاده شده است.

#### ۴- اعمال الگوریتم‌های پیشنهادی

در این تحقیق الگوریتم طبقه‌بندی  $K$ -means نیمه نظارت شده با سه مرحله بازخورد مرتبط پیشنهاد می‌شود. الگوریتم  $K$ -means به دلیل پیچیدگی زمانی  $O(n)$  در کاربردهایی که در تعامل با کاربر هستند بسیار مناسب است. اما مشکل اصلی الگوریتم  $K$ -means همگرایی به بهینه محلی در صورت انتخاب نامناسب مراکز اولیه خوشه است. برای حل این مشکل از تعداد کمی داده آموزشی در هر خوشه برای پیدا کردن محدوده خوشه به صورت نظارت شده استفاده می‌کنیم. در بسیاری از شرایط پیدا کردن محدوده خوشه کافی نیست. چون تنوع داده‌ها زیاد است و استفاده از ویژگی‌های سطح پایین به تنهایی نمی‌تواند به طبقه‌بندی دقیقی دست یافت و برخی از تصاویر اشتباهاً در خوشه‌های همسایه قرار می‌گیرند. برای حل این مشکلات از بازخورد مرتبط ترکیبی استفاده می‌شود. ابتدا ویژگی‌های سطح پایین از تصاویر پایگاه داده استخراج می‌شود. تصاویر پایگاه داده توسط معیار شباهت ماهالانوبیس که در زیر ارائه شده است طبقه‌بندی می‌شوند.  $\sum$  ماتریس کوواریانس کل پایگاه داده است که عناصر غیر

قطری آن صفر شده است. بعد از طبقه‌بندی، مقدار عضویت فازی هر تصویر به خوشه‌ها بر اساس فاصله تصویر تا خوشه محاسبه می‌شود. در مرحله آخر با ارائه یک روش بازخورد مرتبط ترکیبی جدید برای طبقه‌بندی سعی در بهبود دقت سیستم است. تصاویری که کاربر نامربوط تشخیص داده است تصاویری است که متعلق به خوشه‌های همسایه است و سیستم به اشتباه تصویر را به این خوشه تخصیص داده است. برای طبقه‌بندی درست این تصاویر از بازخورد مرتبط طولانی‌مدت استفاده می‌شود. بنابراین مقدار تعلق فازی به دست آمده بعد از مرحله طبقه‌بندی را برای تصویر نامربوط به خوشه را صفر قرار داده و آن را به خوشه‌ای که بیشترین درجه عضویت دارد تخصیص داده می‌شود. آزمایش‌ها نشان می‌دهند که تصاویر نامربوط به‌طور میانگین بعد از دو بازخورد در خوشه درست قرار می‌گیرند. بعد از هر مرحله بازخورد با استفاده از تکنیک تغییر مرکز خوشه (CCM<sup>1</sup>) که مبتنی بر روش تغییر نقطه پرس‌وجو است برای تعیین نقطه ایده‌آل مرکز خوشه در فضای ویژگی با استفاده از رابطه زیر استفاده می‌شود.

$$\mathbf{X}_i^{(j+1)} = \alpha \mathbf{X}_i^{(j)} + \beta \sum_{Y_k \in R} \frac{Y_k}{|R|} - \gamma \sum_{Y_k \in N} \frac{Y_k}{|N|} \quad (8)$$

$\alpha$ ،  $\beta$  و  $\gamma$  ضرایب تنظیم هستند و  $\alpha$  را برابر  $|C - R - N| * 3$  و  $\beta$  و  $\gamma$  را برابر  $\frac{1}{3}$  در نظر می‌گیریم. بدین معنی است که کاربر  $i$  امین خوشه پایگاه داده را بعنوان پرس‌وجو داده‌است و در تکرار  $j$ ام RF کاربر  $X_i^{(j)}$  فرمول‌بندی جاری است که در فرمول‌بندی مرحله بعد به  $X_i^{(j+1)}$  تبدیل می‌گردد و  $R$  کل تصاویری است که کاربران مربوطی به این خوشه دانسته‌اند و  $N$  تصاویر نامربوط این جلسه پرس‌وجو است  $C$  تعداد کل تصاویر متعلق به این خوشه‌است و  $Y_k$  تصاویر متعلق به ناحیه  $N$  و  $R$  است. بازخورد مرتبط تا زمانی که کاربر از نتایج طبقه‌بندی به دست آمده راضی شود ادامه می‌یابد. چون شکل دارای اهمیت بالاتری است و توسط کاربر دیگر متفاوت در نظر گرفته شود چون به رنگ اهمیت بیشتری داده است. بدین منظور برای تخمین وزن ویژگی هر خوشه از تصاویری که کاربر مربوط شناخته است استفاده می‌شود. چون اهمیت ویژگی‌های مختلف در هر خوشه متفاوت است، از روش بازخورد مرتبط Re-weighting برای به دست آوردن وزن ویژگی مناسب هر خوشه بر اساس پراکندگی ویژگی‌های آنها استفاده می‌شود. پس از به دست آمدن وزن ویژگی هر خوشه با ارائه یک معیار شباهت چندگانه به‌صورت رابطه زیر سعی در تعیین فاصله مناسب تصویر پرس‌وجو با خوشه‌مورد نظر است [7].

$$\text{cluster}(I) = \arg \min_{j(1..k)} \sqrt{(\mathbf{F}_I - \mathbf{C}_j)^T \mathbf{W}_j (\mathbf{F}_I - \mathbf{C}_j)} \quad (9)$$

در پاسخ به تصویر پرس‌وجوی  $I$ ، خوشه  $J$  ام که عبارت (9) را مینیمم کند را به‌عنوان پاسخ برگردانده می‌شود که  $\mathbf{W}_j$  بردار وزنی اختصاص یافته به هر خوشه است که در مرحله طبقه‌بندی از یک بردار ثابت که برابر معکوس واریانس ویژگی‌های کل پایگاه داده بود استفاده می‌شد که با اختصاصی شدن آن فاصله‌ها را به مراتب دقیق‌تر محاسبه می‌کند. برای انجام آزمایش‌های لازم جهت ارزیابی سیستم ابتدا به معرفی پایگاه داده‌های مورد استفاده و ویژگی‌های مورد استفاده پرداخته و در بخش بعد آزمایش‌های لازم روی سیستم پیشنهادی را ارائه می‌شود. پایگاه داده مورد استفاده ما شامل ۱۰۰۰ تصویر از مجموعه corel در ده گروه در ده گروه (خوشه‌ها بترتیب ۱- آفریقا ۲- ساحل ۳- توبوس ۴- بنای تاریخی ۵- دایناسور ۶- فیل ۷- گل ۸- اسب ۹- غذا ۱۰- کوهستان است) در ابعاد 256×384 استفاده شده است. ویژگی‌های سطح پایین مورد استفاده ما شامل ۹ ویژگی میانگین، واریانس، کجی‌نمایی از فضای رنگی CIE LUV به‌عنوان ویژگی رنگ است و ۶ ویژگی انرژی، آنتروپی، کنتراست، واریانس، هبستگی و گشتاور تفاضلی معکوس استخراج شده از ماتریس هم‌اتفاقی به‌عنوان ویژگی بافت است و ۱۰ ویژگی از ضرایب زرنیک هم به‌عنوان ویژگی شکل در نظر گرفته می‌شود. برای بررسی تاثیر ترکیب سه روش بازخورد مرتبط طولانی مدت، تغییر وزن ویژگی، تغییر مرکز خوشه و روش ترکیبی با دقت خوشه‌های مختلف طی سه بازخورد در شکل ۳- الف نشان داده شده است.

همانطور که در شکل مشاهده می‌کنید، خوشه‌های ۱، ۳، ۶، ۸ در بعضی بازخوردها دقت طبقه‌بندی کاهش یافته است. بر اساس آزمایش‌ها انجام شده و استفاده از افراد مختلف با انجام چنین بار آزمایش به این نتیجه رسیدیم که حدود ۳ تا ۹ درصد از بازخوردهای کاربران اشتباه است. در حالت کلی سه دلیل عمده برای این مشکل می‌توان ارائه داد. اول اینکه انسان احساسی و غیرقابل پیش‌بینی است و قضاوت‌های او بسیار ذهنی است. دوم اینکه تصاویر واقع در نواحی مرزی خوشه‌ها بسیار شبیه هم

<sup>1</sup> Cluster center movement

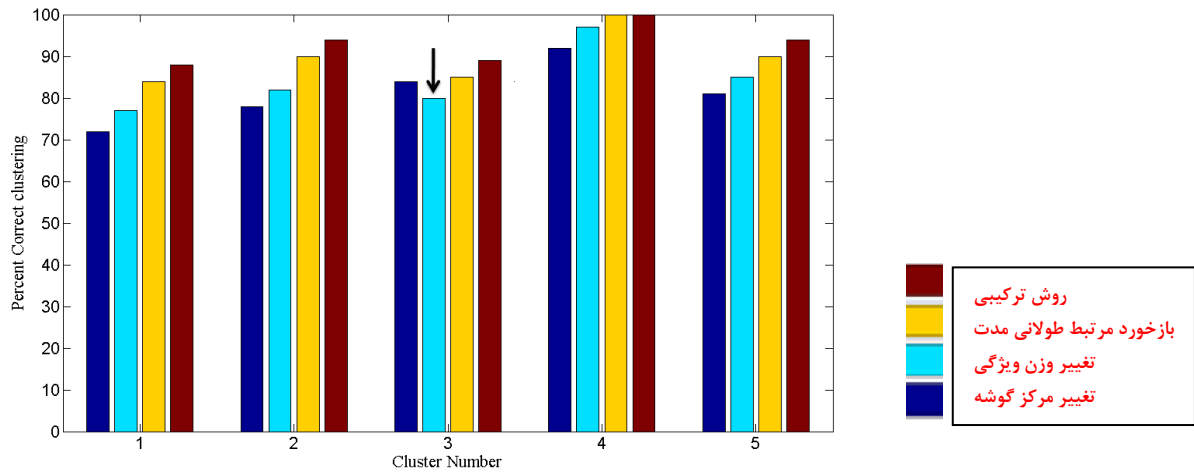


هستند و کاربر ممکن است اشتباهها آن را مربوط تشخیص دهد که باعث کاهش دقت طبقه‌بندی می‌شود. سوم اینکه کاربر در برخی موارد ممکن است کاربر بخواهد تصویری را نامربوط علامت‌گذاری کند ولی آن را اشتباه مربوط علامت می‌زند و یا برعکس. در مورد تصاویر ماهواره‌ای نیز این مشکل وجود دارد ولی به مراتب کمتر و تنها در خوشه سه این کاهش دقت وجود دارد. دلیل این امر نیز وجود تفاوت‌های بارز بین تصاویر خوشه‌های مختلف است و دلایل اول و دوم در اینجا کمتر احساس می‌شود.

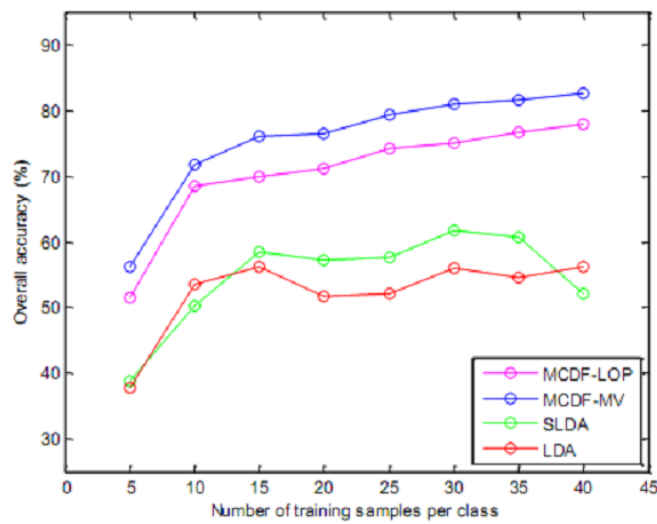
در شکل ۳- ب نیز می‌توانید نمودار دقت را طی بازخوردهای مختلف برای سه روش بازخورد مرتبط طولانی مدت، تغییر وزن ویژگی، تغییر مرکز خوشه و روش ترکیبی در مقایسه با روش دیگر را ببینید. همانطور که ملاحظه می‌شود افزایش دقت طبقه‌بندی به‌طور میانگین در روش ترکیبی حدود ۲۰٪ است در صورتی هر سه روش به‌طور جداگانه افزایش دقتی حدود ۹-۱۱٪ را دارد. در روش پیشنهادی با ترکیب سه روش پیشنهادی حداکثر استفاده از نظرات کاربر برده شده است. سیستم بازیابی تصویر به دلیل وجود فاصله معنایی، ویژگی‌های سطح پایین نمی‌توانند بخوبی تصاویر را طبقه‌بندی کنند. ما از طبقه‌بندی نیمه‌نظارت شده و بازخورد مرتبط برای کاهش شکاف معنایی استفاده نمودیم. برای بهبود طبقه‌بندی مبتنی بر طبقه‌بندی ترکیبی از سه بازخورد مرتبط، طولانی مدت و تغییر وزن ویژگی و تغییر مراکز خوشه استفاده کرده‌ایم. نتایج آزمایش‌ها نیز کارایی روش ترکیبی را در مقایسه با تک تک روش‌ها نشان می‌دهد. روش بازخورد مرتبط ترکیبی روی همه روش‌های طبقه‌بندی مبتنی بر مرکز می‌تواند باعث بهبود سیستم شود.

داده‌های تصویری که در این پژوهش استفاده شده است، شامل ۱۰۰۰ تصویر است. در همه الگوریتم‌ها از طبقه‌بند k-NN استفاده شده است، چون این طبقه‌بند در همه الگوریتم‌ها پاسخ بهتری در نرخ بازشناسی نسبت به شبکه‌های عصبی احتمالاتی داشته است. طبقه‌بندهایی که در این تحقیق استفاده شده است، شامل طبقه‌بند نزدیکترین همسایه و شبکه‌های عصبی احتمالاتی است که با توجه به نتایج اعمالی بر دو پایگاه داده، نشان از برتری طبقه‌بند نزدیکترین همسایه است. با اعمال تبدیل رادون در این داده‌ها و با آنالیز مؤلفه اساسی پاسخ بهتری مشاهده می‌شود، بنابراین از درج مقدار بدون PCA صرف‌نظر می‌شود. محدوده زاویه نیز ۰ تا ۳۶۰ درجه در نظر گرفته شده است و معیار فاصله استفاده شده در طبقه‌بند در همه الگوریتم‌ها از نتایج قسمت‌های قبلی برگرفته شده است.

در الگوریتم پیشنهادی ناحیه‌بندی تصاویر از بلوک  $16 \times 16$  استفاده شده است. فواصل میان نقاط کانتور مرزی و مرکز ثقل از تعداد نقطه ۳۰۰ و معیار فاصله میان این نقاط و نقطه مرکز ثقل، فاصله اقلیدسی است. شکل ۴ دقت بازیابی به تعداد عکس بازیابی مختلف به ازای دو طبقه‌بند مختلف نیمه نظارتی PNN و K-NN را نشان می‌دهد. در الگوریتم هیستوگرام هرمی از گرادیان‌های جهت‌دار که تعداد آن برابر با پنج و در سه مرحله استفاده شده است. همچنین محدوده زاویه را ۰ تا ۳۶۰ درجه در نظر گرفتیم. با توجه به شکل نرخ بازشناسی در همه الگوریتم‌ها با استفاده از داده‌های تصویری با افزایش اندکی همراه است و علت آن زیاد بودن تعداد داده‌های پایگاه داده نسبت به داده است. نرخ بازشناسی بر داده‌های تصویری استفاده‌شده در این تحقیق که الگوریتم هیستوگرام هرمی گرادیان‌های جهت‌دار بر آن‌ها اعمال شده، دارای بالاترین نرخ نسبت به دیگر روش‌های پیشنهادی بوده و ۹۸/۵۲٪ است. نتایج بازیابی در شکل ۵ برای بعضی از گروه‌ها نشان داده شده است. پیچیده‌ترین گروه از بابت شلوغ بودن پس زمینه گروه آفریقای‌ها است که نتیجه بازیابی سیستم در شکل ۶ آورده شده است.

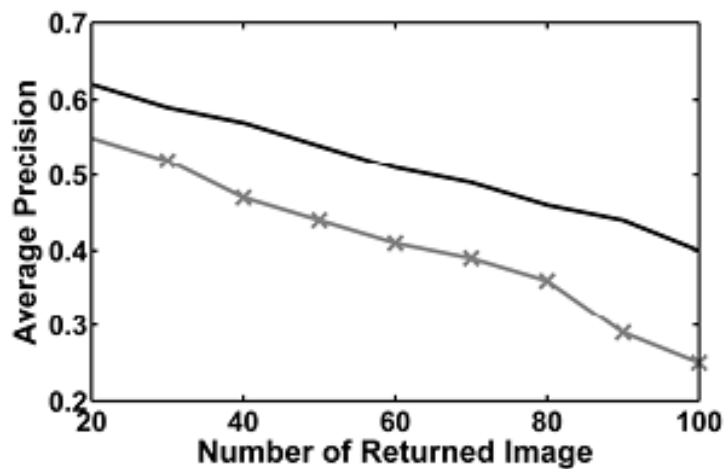


(الف)



(ب)

شکل ۳: الف: تاثیر بازخورد مرتبط طی سه بازخورد در دقت طبقه‌بندی روی پایگاه داده Corel ب: دقت بازیابی به ازای چند روش مختلف



شکل ۴: دقت بازیابی به تعداد عکس بازیابی مختلف به ازای دو طبقه بند مختلف نیمه نظارتی



(الف)



(ب)



(ج)



(د)

شکل ۵: الف) ۹ تصویر صحیح از ۱۰ تصویر ب) ۱۰ تصویر صحیح از ۱۰ تصویر ج) ۸ تصویر صحیح از ۱۰ تصویر د) ۵ تصویر صحیح از ۱۰ تصویر



شکل ۶: پنج تصویر برگردانده شده توسط سیستم پیشنهادی در گروه آفریقاییان

## ۵- نتیجه گیری

سیستم‌های نیمه‌نظارت شده و بدون ناظر به دلیل کاهش کارهای کاربر امروزه بیشتر مورد توجه قرار گرفته‌اند. بخصوص در زمینه‌هایی مثل بازیابی تصاویر روی وب که تعداد تصاویر زیاد است و گرفتن نظرات کاربر بسیار وقت‌گیر است، بنابراین همواره استفاده از نظرات چندین کاربر کارایی بسیار بالاتری دارد. در این مقاله بعد از پیش‌پردازش و تهیه داده‌های تصویری، الگوریتمی برای استخراج ویژگی پیشنهاد شده است. حساس نبودن این الگوریتم نسبت به چرخش از مزیت عمده آن است، زیرا که نمونه‌برداری نقاط از بیشترین فاصله از مرکز ثقل شروع شده و با چرخش تصاویر در هر کدام از حالت‌ها بردار ویژگی مربوط به تصویر ثابت می‌ماند. از دیگر مزیت عمده این الگوریتم، کم بودن زمان پردازش آن است. در این پژوهش از دو روش طبقه‌بندی حالات استفاده شده است که شامل طبقه‌بند  $k$ -NN و PNN است که با توجه به نتایج در همه روش‌های پیشنهاد شده، از طبقه‌بند  $k$ -NN پاسخ بهتری نسبت به PNN مشاهده شده است. همچنین افزایش دقت طبقه‌بندی به‌طور میانگین در روش ترکیبی حدود ۲۰٪ است و نیز می‌توان به این نتیجه رسید که سیستم پیشنهادی در این پژوهش توانسته یک طبقه‌بند نیمه‌نظارتی به همراه کاهش فاصله معنایی را در بازیابی ارائه کند.

## مراجع

- [1] F. Cao , J. Liang and B. Liang, “ A new initialization method for categorical data clustering,” *Expert Systems with Applications.*, vol 36,no.7, pp. 10223–10228, 2009 ,doi:10.1016/j.eswa.2009.01.060.
- [2] A. Amato and V. Lecce, “ A knowledge based approach for a fast image retrieval system,” *Image and Vision Computing* vol 26,no.11, pp. 1466–1480, 2008, doi:10.1016/j.imavis.2008.01.005.
- [3] Y. Chen, J. Z. Wang and R. Krovetz, "CLUE: cluster-based retrieval of images by unsupervised learning," in *IEEE Transactions on Image Processing*, vol. 14, no. 8, pp. 1187-1201, Aug. 2005, doi: 10.1109/TIP.2005.849770.
- [4] R. Zhang and Zh. Zhang, “ Empirical Bayesian learning in the relevance feedback for image retrieval,” *Image and Vision Computing*. vol. 24,no.3, pp.211–223, 2006, doi:10.1016/j.imavis.2005.11.004.
- [5] V.Mezaris and I. Kompatsiaris, “ Region Based Image Retrieval Using an Object Ontology and Relevance Feedback,” *EURASIP Journal on Applied Signal Processing* ,vol.6,pp. 886–901, 2004, doi: 10.1155/S1110865704401188
- [6] Y. Rui, T. S. Huang and S. Mehrotra, "Content-based image retrieval with relevance feedback in MARS," *Proceedings of International Conference on Image Processing*, 1997, pp. 815-818 vol.2, doi: 10.1109/ICIP.1997.638621.
- [7] J.A. Hartigan and M.A. Wong , “Algorithm AS136: A k-means Clustering algorithm,” *Applied Statistic*, vol.28,no.1, pp. 100-108. 1979, doi:10.2307/2346830.
- [8] B. Kimia, “Shape Representation for Image Retrieval”, *Image Databases: Search and Retrieval of Digital Imagery*, John Wiley & Sons, pp. 345-358, 2001, doi:10.1002/0471224634.ch13.

- [9] J.Li, J. Z. Wang and G .Wiederhold , “Integrated Region Matching for Image Retrieval,” *ACM Multimedia*, p. 147-156,2000 ,doi: 10.1145/354384.354452,
- [10] J. Mao and A.K. Jain, “Texture Classification and Segmentation using Multi-Resolution Simultaneous Autoregressive Models,” *Pattern Recognition*, vol. 25,no.2, pp. 173-188, 2010, doi:10.1016/0031-3203(92)90099-5.
- [11] M. Jalali and T. Sedghi, "Semi Supervised Feature Extraction for Filling Semantic Gap in Image Retrieval," *Iranian Conference on Machine Vision and Image Processing*, 2011, pp. 1-4, doi: 10.1109/IranianMVIP.2011.6121537.
- [12] A. Pentland, R. Picard and S. Sclaroff “Photobook: Contentbased manipulation of image databases,” *International Journal of Computer Vision*, vol.18,no.3, pp.233–254 ,1996, doi:10.1007/BF00123143.
- [13] T. N. Kipf and M. Welling, “Semi-supervised classification with graph convolutional networks,” *5th Int. Conf. Learn. Represent. ICLR 2017 - Conf. Track Proc.*, pp. 1–14, 2017.
- [14] R. Zhang and Z. Zhang " BALAS: Empirical Bayesian learning in the relevance feedback for image retrieval," *Image and Vision Computing* , vol.24, no.3, pp. 211–223,2006, doi:10.1016/j.imavis.2005.11.004.
- [15] J. Smith, “*Color for Image Retrieval*”, Image Databases: Search and Retrieval of Digital Imagery, John Wiley & Sons, New York, pp.285-311,2001.
- [16] J. R. Smith and Shih-Fu Chang, "Single color extraction and image query," *Proceedings., International Conference on Image Processing*, 1995, pp. 528-531 vol.3, doi: 10.1109/ICIP.1995.537688.
- [17] M. Stricker and M. Swain, "The capacity of color histogram indexing," *1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1994, pp. 704-708, doi: 10.1109/CVPR.1994.323774.
- [18] M. Unser, "Texture classification and segmentation using wavelet frames," in *IEEE Transactions on Image Processing*, vol. 4, no. 11, pp. 1549-1560, Nov. 1995, doi: 10.1109/83.469936.
- [19] K. P. Yip, D. W. Cheung and M. K. Ng, "On discovery of extremely low-dimensional clusters using semi-supervised projected clustering," *21st International Conference on Data Engineering (ICDE'05)*, 2005, pp. 329-340, doi: 10.1109/ICDE.2005.96.
- [20] L. Nanni, A. Rigo, A. Lumini, and S. Brahmam, “Spectrogram classification using dissimilarity space,” *Appl. Sci.*, vol. 10, no. 12, pp. 1–17, 2020, doi:10.3390/app10124176.
- [21] Y. Chen, J. Z. Wang and R. Krovetz, "CLUE: cluster-based retrieval of images by unsupervised learning," in *IEEE Transactions on Image Processing*, vol. 14, no. 8, pp. 1187-1201, Aug. 2005, doi: 10.1109/TIP.2005.849770.
- [22] K. Z. Mao, K. -C. Tan and W. Ser, "Probabilistic neural-network structure determination for pattern classification," in *IEEE Transactions on Neural Networks*, vol. 11, no. 4, pp. 1009-1016, July 2000, doi: 10.1109/72.857781.
- [23] M. Flickner *et al.*, "Query by image and video content: the QBIC system," in *Computer*, vol. 28, no. 9, pp. 23-32, Sept. 1995, doi: 10.1109/2.410146.
- [24] X. Zhang, M. Lei, D. Yang, Y. Wang and L. Ma, “Multi-scale curvature product for robust image corner detection in curvature scale space,” *Pattern Recognition Letters*, vol.28,no.5,Apr. 2007, doi:10.1016/j.patrec.2006.10.006.
- [25] N. C. Shirazi, R. Hamzehyan, and A. Masoomi, " The Comparison of Classification Algorithms for Remote Sensing Images," *Journal of Communication Engineering.*, vol. 5,no.17, pp. 31-38, 2015(in Persian).

- [26] A. Soltani-Farani, H. R. Rabiee, and S. A. Hosseini, "Spatial-aware dictionary learning for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 1, pp. 527-541, 2015, doi: 10.1109/TGRS.2014.2325067.