

بازیابی معنایی تصویر بر اساس ترکیب روش‌های یادگیری کوتاه‌مدت

بهاره باقری، مریم پورمحمی‌آبادی و حسین نظام‌آبادی پور

یادگیری با نظارت است که روی ارتباط بین کاربر و موتور جستجو تمرکز می‌کند و در این فرایند کاربر بر اساس پرس و جوی مد نظر خود تصاویر بازیابی شده توسط سامانه را با عناوین مرتبط یا غیر مرتبط، برچسب‌گذاری می‌کند. کاربر و سامانه تا رسیدن به تصاویر دلخواه کاربر با یکدیگر تعامل دارند. یادگیری با استفاده از بازخورد ربط به دو صورت انجام می‌گیرد: یادگیری کوتاه‌مدت و یادگیری بلندمدت [۱] و [۲].

یادگیری کوتاه‌مدت (STL) در ارتباط با یک جستجوی کاربر و با هدف هدایت سامانه به سمت خواسته مطلوب کاربر انجام می‌شود و طی آن، یادگیری در خلال یک پرس و جوی خاص صورت می‌گیرد. در تعامل هر کاربر با سامانه، سامانه به کمک یادگیری کوتاه‌مدت سعی در کشف معنای مورد نظر کاربر می‌کند. روش‌های یادگیری کوتاه‌مدت را می‌توان به چهار دسته طبقه‌بندی کرد: روش‌های مبتنی بر طبقه‌بندی، روش‌های مبتنی بر بهبود تابع شباهت، روش‌های مبتنی بر بهبود بردار پرس و جوی و روش‌های چندپرسی [۳] و [۴].

در روش‌های مبتنی بر طبقه‌بندی، یک طبقه‌بند برای دسته‌بندی تصاویر به دو گروه مرتبط و نامرتب با استفاده از بردار ویژگی تصاویر، طراحی می‌شود [۲]، [۵] و [۶].

در روش بهبود تابع شباهت، وزن ویژگی‌ها در یک تابع شباهت در خلال بازخورد ربط به منظور رسیدن به هدف کاربر، تغییر می‌کند [۴]، [۷] و [۸].

در روش‌های مبتنی بر بهبود بردار پرس و جوی، این بردار از تصاویر نامرتب دور و به تصاویر مرتبط نزدیک می‌شود [۹] و [۱۰].

در روش‌های چندپرسی، از روش‌های مختلف برای تهیه بردارهای پرس و جوی بیشتر استفاده می‌شود [۱۱] تا [۱۳].

بسیاری از الگوریتم‌های یادگیری در حقیقت یک نوع جستجوی محلی انجام می‌دهند که ممکن است در کمینه‌های محلی گرفتار شوند. همچنین هر روش یادگیری دارای نقاط ضعف و قوت متفاوتی است و با توجه به مقدار پارامترهایش، به پاسخ متفاوتی برای مسأله می‌رسد. در نتیجه ترکیب اطلاعات از چند روش متفاوت نتایج بهتری را فراهم می‌کند و اجازه توزیع طبیعی و انعطاف‌پذیری از منابع را می‌دهد. به عبارت دیگر، هدفش این است که کارایی سامانه بازیابی تصویر را افزایش داده و نقاط ضعف یک روش را با نقاط قوت روش دیگر پوشش دهد. در راهکار ترکیب، روش‌های یادگیری در یک ساختار موازی قرار دارند. الگوی ورودی به همه روش‌های یادگیری اعمال می‌شود و هر کدام از آنها یک جواب برای مسأله ارائه می‌کنند. سپس با ترکیب مناسب این جواب‌ها پاسخ نهایی به دست می‌آید [۱۴] تا [۱۶].

در این مقاله، روشی جدید در بازخورد ربط ارائه می‌شود که از ترکیب چند روش یادگیری کوتاه‌مدت استفاده می‌کند. نوآوری این مقاله در ارائه یک راه‌یافت برای استفاده از دو یا چند روش یادگیری کوتاه‌مدت و نحوه ترکیب اطلاعات ناشی از آنها در بازیابی تصویر است. در روش پیشنهادی، ترکیب در سه سطح انجام می‌گیرد: سطح تصاویر برگردانده شده، سطح

چکیده: بازیابی تصویر بر اساس محتوا شامل مجموعه‌ای از روش‌ها برای پردازش ویژگی‌های دیداری یک تصویر پرس و جو به منظور پیدا کردن تصاویر مشابه آن در یک پایگاه تصویر است. برای بهبود عملکرد سامانه بازیابی تصویر می‌توان از ابزار بازخورد ربط استفاده کرد. در این تحقیق برای افزایش کارایی سامانه‌های بازیابی تصویر، ترکیب دو یا چند روش یادگیری کوتاه‌مدت در فرایند بازخورد ربط پیشنهاد شده است. در روش پیشنهادی، ترکیب در سه سطح انجام می‌گیرد: سطح تصاویر برگردانده شده، سطح توابع شباهت و سطح رتبه تصاویر. برای ارزیابی روش پیشنهادی یک سامانه بازیابی تصویر مبتنی بر محتوا با ۱۰۰۰۰ تصویر از ۸۲ گروه معنایی متفاوت مورد استفاده قرار گرفته است. نتایج حاصل از آزمایش‌ها برتری روش پیشنهادی را از لحاظ دقت بازیابی مورد تأیید قرار می‌دهد.

کلیدواژه: بازیابی تصویر مبتنی بر محتوا، بازخورد ربط، شکاف معنایی، ترکیب، یادگیری کوتاه‌مدت.

۱- مقدمه

با پیشرفت‌های به وجود آمده در فناوری اطلاعات، رشد انفجارگونه‌ای از پایگاه‌های تصویر ایجاد شده که نیاز کاربران به ابزارهای مؤثر و مفید برای جستجو در این مجموعه‌های بزرگ را افزایش می‌دهد. به طور کلی، دو رویکرد برای جستجو در مجموعه‌های تصویری وجود دارد: روش مبتنی بر متن و روش مبتنی بر محتوا. در اولین رویکرد، تصاویر با استفاده از کلمات کلیدی شرح داده می‌شوند و برای بازیابی تصاویر از کلیدواژه‌ها استفاده می‌شود. این روش محدودیت‌هایی دارد که عبارتند از: الف) برچسب‌گذاری تصاویر پایگاه داده، کاری دشوار و زمان‌بر است و ب) برچسب‌گذاری به صورت اصولی انجام نمی‌گیرد و کاربران متفاوت از کلمات مختلفی برای شرح یک تصویر استفاده می‌کنند.

برای غلبه بر این مشکلات، سیستم‌های بازیابی تصویر مبتنی بر محتوا مورد توجه قرار گرفتند. در این سیستم‌ها از الگوریتم‌های پردازش تصویر برای استخراج خودکار ویژگی‌هایی چون رنگ، بافت، شکل و ... از تصاویر استفاده می‌شود. همچنین این سیستم تصاویری را بازیابی می‌کند که از لحاظ معنا، مشابه با تصویر پرس و جو باشند. مهم‌ترین چالش فراروی این رویکرد عبارت است از وجود شکاف معنایی^۱ میان ویژگی‌های بصری سطح پایین و معنای سطح بالای موجود در تصاویر.

بازخورد ربط روش مؤثری برای پرکردن این شکاف و بهبود عملکرد سامانه‌های بازیابی به شمار می‌آید. در حقیقت، بازخورد ربط یک روش

این مقاله در تاریخ ۱۰ آذر ماه ۱۳۹۲ دریافت و در تاریخ ۱۲ آبان ماه ۱۳۹۳ بازنگری شد.

بهاره باقری، بخش مهندسی برق، دانشکده فنی و مهندسی، دانشگاه شهید باهنر کرمان، کرمان، (email: bagheri_bhr@yahoo.com).

مریم پورمحمی‌آبادی، بخش مهندسی برق، دانشکده فنی و مهندسی، دانشگاه شهید باهنر کرمان، کرمان، (email: pmahyabadi@iust.ac.ir).

حسین نظام‌آبادی پور، بخش مهندسی برق، دانشکده فنی و مهندسی، دانشگاه شهید باهنر کرمان، کرمان، (email: nezam@mail.uk.ac.ir).

$$Q = \frac{1}{|Q^+|} \sum_{q^+ \in Q^+} q^+ \quad (3)$$

در (۲) تابع هزینه J_{R-N} به صورت نسبت فاصله دورترین بردار ویژگی مرتبط از مجموعه Q^+ به نزدیکترین بردار ویژگی غیر مرتبط از مجموعه Q^- تا بردار پرس و جوی Q تعریف شده است. کمینه کردن این تابع متناظر با حداکثر کردن حاشیه طبقه بندی در دو دسته تصاویر مرتبط و غیر مرتبط می باشد. جهت حداقل کردن این تابع از روش کمینه سازی به روش گرادینان نزولی استفاده شده و در نتیجه لازم است مشتق های این تابع نسبت به متغیر w محاسبه شوند.

روش دیگری در این دسته، از اطلاعات رتبه بندی تصاویر مرتبط برای به روز رسانی وزن های هر ویژگی استفاده می کند [۷]. در این روش برای تنظیم وزن ها، بازیابی تصویر بر اساس تک تک ویژگی ها، بدون در نظر گرفتن سایر ویژگی ها انجام می شود. سپس رتبه تصاویری که به عنوان مرتبط مشخص شده اند، در بازیابی با هر یک از ویژگی ها محاسبه می شود. پس از آن وزن مربوط به هر ویژگی متناسب با عکس مجموع رتبه های تصاویر مرتبط برای آن ویژگی در نظر گرفته می شود. همچنین برای تنظیم وزن های مؤلفه های ویژگی که یک روش قاعده- پایه است از اطلاعات واریانس و میانگین هر مؤلفه روی تصاویر مرتبط و نامرتب استفاده کرده است.

۲-۲ روش یادگیری مبتنی بر طبقه بندی

در این نوع روش ها از یک طبقه بند برای جداسازی تصاویر مرتبط و نامرتب استفاده می شود. ماشین بردار پشتیبان $(SVM)^2$ ، طبقه بند بیز^۳، درخت تصمیم^۴ و شبکه عصبی^۵ از جمله روش های طبقه بندی هستند که برای یادگیری مفاهیم سطح بالا از ویژگی های سطح پایین تصاویر به کار می روند. از بین این طبقه بندها، SVM به علت داشتن قدرت تعمیم پذیری بالا و توانایی طبقه بندی داده های پراکنده در یک فضای با بعد بالا، از عملکرد بسیار خوبی برخوردار است و به وفور در تحقیقات استفاده می شود. زمانی که تعداد نمونه های مرتبط کم باشد، عملکرد طبقه بند SVM به دلیل ناپایداری، مشکل بیش برآزش و گرایش به سمت داده های نامرتب، ضعیف می شود. در [۱۸] و [۱۹] برای رفع مشکلات موجود از ماشین بردار پشتیبان با یادگیری فعال استفاده شده است. در یادگیری فعال، SVM دورترین نمونه ها از ابرصفحه که در سمت تصاویر مرتبط واقع شده اند را به عنوان مرتبط ترین تصاویر در نظر می گیرد و نزدیک ترین نمونه ها به ابرصفحه را به عنوان تأثیرگذارترین تصاویر در نظر گرفته و برای برچسب گذاری به کاربر نمایش می دهد.

۲-۳ روش یادگیری مبتنی بر بهبود بردار پرس و جو

اولین رابطه بهبود بردار پرس و جو در زمینه بازیابی تصویر در سال ۱۹۷۰ توسط روچو^۶ بیان شد [۲۰]. این روش سعی بر این دارد که بردار پرس و جو را به سمت داده های مرتبط حرکت دهد و از داده های نامرتب دور کند. همچنین این روش اولین بار در سال ۱۹۹۷ در سامانه بازیابی محتوایی تصویر MARS پیشنهاد شد [۲۱]. در این روش فرض بر این

توابع شباهت و سطح رتبه تصاویر.

در بخش دوم به بررسی اجمالی تحقیقات انجام شده در زمینه یادگیری کوتاه مدت می پردازیم و چالش های موجود در این زمینه را بیان خواهیم کرد. در بخش سوم روش پیشنهادی معرفی شده و در بخش چهارم نتایج عملکرد سیستم بازیابی و مقایسه با سایر روش ها ارائه می شود. در نهایت در بخش پنجم، نتیجه گیری و جمع بندی مقاله می آید.

۲- تحقیقات مرتبط

از اوایل ۱۹۹۰، بازیابی تصویر بر اساس محتوا، یک زمینه فعال برای تحقیقات قلمداد می شود. تا به حال سامانه های بازیابی تجاری و تحقیقاتی فراوانی ایجاد و معرفی شده است. بازخورد ربط یکی از روش های ارائه شده برای کاهش فاصله معنایی محسوب می شود که در سال ۱۹۹۸ برای اولین بار توسط روی^۱ به کار گرفته شد [۱۷]. در سامانه بازیابی تصویر مبتنی بر محتوا، در هر مرحله از بازیابی، کاربر با سامانه ارتباط برقرار کرده و سامانه با استفاده از روش یادگیری کوتاه مدت، معنای مطلوب و مورد نظر کاربر را یاد می گیرد. روش های یادگیری کوتاه مدت را می توان به چهار دسته طبقه بندی کرد: روش های مبتنی بر طبقه بندی، روش های مبتنی بر بهبود تابع شباهت، روش های مبتنی بر بهبود بردار پرس و جو و روش های چندپرسی [۳] و [۴].

۲-۱ روش یادگیری مبتنی بر بهبود تابع شباهت

در روش بهبود تابع شباهت، وزن های تابع شباهت در خلال بازخوردها تنظیم می شوند. برخی از روش های موجود در این دسته، وزن مربوط به ویژگی ها را با استفاده از اطلاعات میانگین و واریانس محاسبه می کنند. برخی دیگر تابع خطای مناسبی را تعریف می کنند و وزن های مربوط به تابع شباهت را با کمینه کردن تابع خطا از طریق الگوریتم گرادینان نزولی، یاد می گیرند. استفاده از گرادینان نزولی برای بهبود طبقه بندی در [۴] گزارش شده است. در این تحقیق، تابع شباهت به صورت (۱) تعریف می شود

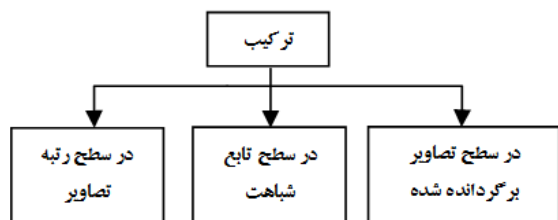
$$d^r(Q, F|W) = w_{1,j}^r \sum_{j=1}^k (\frac{q_{1,j} - f_{1,j}}{q_{1,j} + f_{1,j}})^r + \dots + w_{k,j}^r \sum_{j=1}^k (\frac{q_{k,j} - f_{k,j}}{q_{k,j} + f_{k,j}})^r \quad (1)$$

که در آن F و Q بردارهای ویژگی دو تصویر هستند و d شباهت بین دو تصویر را محاسبه می کند. نشان دهنده $f_{k,j}$ نشان دهنده k امین مؤلفه از k امین ویژگی است. l_k نشان دهنده طول بردار k امین ویژگی است، $w_{k,j}$ وزن k امین ویژگی و $w_{k,j|j \neq k}$ وزن k امین مؤلفه از k امین ویژگی است. k تعداد ویژگی ها را مشخص می کند.

تابع خطا به صورت (۲) تعریف می شود. در این رابطه، f^+ دورترین بردار ویژگی مرتبط به تصویر پرس و جوی Q و f^- نزدیکترین بردار ویژگی غیر مرتبط به آن در مجموعه تصاویر بازگردانده شده توسط سامانه هستند. بردار ویژگی در هر مرحله طبق (۳) به صورت میانگین بردار ویژگی های مجموعه تصاویر مرتبط در نظر گرفته شده است [۴]

$$J_{R-N} = \frac{d(Q, F^+|W)}{d(Q, F^-|W)} \quad (2)$$

2. Support Vector Machine
3. Bayesian Classifier
4. Decision Tree
5. Artificial Neural Network
6. Most Informative
7. Rocchio



شکل ۲: روش‌های ترکیب اطلاعات.

روش‌های یادگیری کوتاه‌مدت پایه باید توجه داشت که هر یک عملکرد متفاوتی داشته و نتایج یکسان نداشته باشند تا ترکیب نتایج آنها باعث بهبود نرخ بازشناسی سیستم ترکیبی شود که در این راستا می‌توان از روش‌های یادگیری کوتاه‌مدت مبتنی بر طبقه‌بند، روش‌های بهبود بردار پرس و جو، روش‌های بهبود تابع شباهت و روش‌های چندپرسی استفاده کرد.

نحوه ترکیب خروجی روش‌های یادگیری یا به عبارت دیگر انتخاب قاعده ترکیب به گونه‌ای انتخاب می‌شود که نرخ بازشناسی سیستم ترکیبی بیشینه باشد. قواعد مختلفی برای ترکیب اطلاعات وجود دارد که می‌توان در سه سطح تقسیم‌بندی کرد [۲۴] (شکل ۲).

۳-۱ ترکیب در سطح تصاویر برگردانده‌شده

مجموعه تصاویر برگردانده‌شده به کاربر را با مجموعه R نشان می‌دهیم که شامل r_R تصویر است. در روش ترکیب اطلاعات در سطح تصاویر برگردانده‌شده، این مجموعه به دو قسمت R_{s1} و R_{s2} تقسیم می‌شود. این مجموعه‌ها شامل تصاویر کوتاه‌مدت هستند که به ترتیب توسط روش‌های یادگیری کوتاه‌مدت ۱ و ۲ (STL_1 و STL_2) برگردانده شده‌اند و همچنین داریم $r_R = r_{s1} + r_{s2}$.

برای تعیین تعداد تصاویر هر یک از این دو مجموعه از توابع احتمال P_{s1} و P_{s2} استفاده می‌شود. در شروع هر نشست فرض می‌شود $P_{s1} = P_{s2}$. پس از آن مقادیر P_{s1} و P_{s2} به کمک (۴) محاسبه می‌شوند. در این رابطه $|Q_{s1}^+|$ و $|Q_{s2}^+|$ به ترتیب تعداد تصاویر مرتبطی هستند که توسط هر کدام از روش‌های کوتاه‌مدت در مرحله قبل برگردانده شده‌اند. مقادیر r_{s1} و r_{s2} مطابق (۵) با P_{s1} و P_{s2} نسبت مستقیم دارند، یعنی طبق (۴)، هر کدام از مجموعه‌های R_{s1} و R_{s2} که دارای تصاویر مرتبط بیشتری باشند در تکرار بعد تصاویر بیشتری را برمی‌گرداندند

$$P_{s1} = \frac{|Q_{s1}^+|}{|Q_{s1}^+| + |Q_{s2}^+|} \quad (4)$$

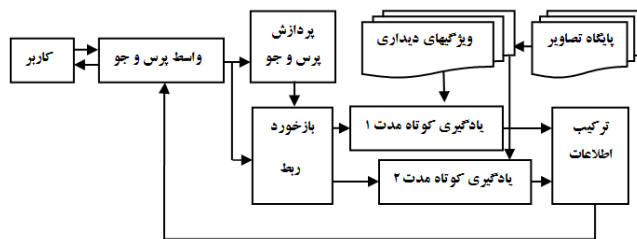
$$P_{s2} = \frac{|Q_{s2}^+|}{|Q_{s1}^+| + |Q_{s2}^+|}$$

$$r_{s1} = r_R \cdot P_{s1} \quad (5)$$

$$r_{s2} = r_R \cdot P_{s2}$$

۳-۲ ترکیب اطلاعات در سطح رتبه تصاویر

در این روش، لیست‌های به دست آمده از رتبه تصاویر در روش‌های یادگیری کوتاه‌مدت با هم ترکیب شده و تنها یک لیست رتبه‌بندی محاسبه می‌شود، سپس لیست نهایی مرتب شده و r_R تصویر با کمترین رتبه انتخاب می‌شوند. در این روش به تصویر با بالاترین میزان شباهت، کمترین رتبه اختصاص داده می‌شود. برای ترکیب در سطح رتبه تصاویر می‌توان از روش‌های وزن‌دار و بدون وزن استفاده کرد.



شکل ۱: ساختار کلی سامانه پیشنهادی جهت بازیابی تصویر با ترکیب اطلاعات حاصل از دو روش یادگیری.

است که تمام تصاویری که از لحاظ معنا مشابه هستند بر اساس برخی اندازه‌گیری‌های فاصله در یک نقطه از فضای ویژگی قرار دارند اما در واقعیت، تصاویر مشابه از لحاظ معنا ممکن است با هم کاملاً متفاوت باشند و در فضای ویژگی به صورت زیرفضاهایی پراکنده شوند.

۲-۴ روش یادگیری چندپرسی

در تبدیل بردار پرس و جو به چند بردار، مشکل موجود در روش بهبود بردار پرس و جو حل خواهد شد و آزمایش‌ها نشان داده‌اند که این روش نسبت به بهبود بردار پرس و جو کارآمدتر است. کاربر چندین تصویر را به عنوان تصاویر مرتبط مشخص می‌کند، سپس سامانه بردارهای ویژگی تصاویر مرتبط را خوشه‌بندی کرده و مراکز خوشه‌ها را به عنوان بردارهای پرس و جو جدید انتخاب می‌کند. در نهایت پایگاه داده با هر کدام از بردارهای پرس و جو جستجو شده و نتایج حاصل از آنها با هم ترکیب می‌شوند. تحقیق [۲۲] از الگوریتم خوشه‌بندی سلسله‌مراتبی برای خوشه‌بندی استفاده کرده است. تعداد ایده‌آل خوشه‌ها به طور خودکار تخمین زده شده‌اند و مراکز خوشه‌ها را به عنوان بردارهای پرس و جو جدید انتخاب می‌کند. بعد از جستجوی پایگاه تصاویر با مرکز هر خوشه، برای هر تصویر پایگاه، کمترین فاصله از مراکز خوشه‌ها به عنوان میزان شباهت تصویر به تصویر پرس و جو در نظر گرفته شده و تصاویر بر این اساس مرتب می‌شوند. نوع معیار استفاده‌شده در اندازه‌گیری شباهت تأثیر مهمی در نتایج دارد که در [۲۳] به آن پرداخته شده است.

۳- روش پیشنهادی

ساختار سامانه بازیابی تصویر پیشنهادی در شکل ۱ به تصویر کشیده شده است. ویژگی‌های مناسبی (مانند رنگ، بافت، شکل و ...) از تصاویر استخراج شده و در پایگاه داده ذخیره می‌شوند. پردازش پرس و جو ویژگی‌های نامبرده را از تصویر پرس و جو استخراج می‌کند. واسط پرس و جو با کاربر در تعامل است به شکلی که اطلاعات لازم را از کاربر جمع‌آوری کرده و نتایج بازیابی را به او برمی‌گرداند. هر روش یادگیری کوتاه‌مدت، تصاویری را بازیابی می‌کند که مشابه با تصویر پرس و جو باشد. سپس نتایج به دست آمده از بازیابی تصاویر در روش‌های یادگیری با هم ترکیب شده و تنها یک لیست نهایی از نتایج محاسبه می‌شود. لیست نهایی مرتب شده و k بهترین تصاویر از طریق رابط گرافیکی به کاربر نمایش داده می‌شوند تا با مقادیر مرتبط و نامرتب برچسب‌گذاری گردند.

همان طور که در شکل ۱ مشاهده می‌شود در روش پیشنهادی بلوکی به نام ترکیب اطلاعات وجود دارد که اطلاعات و تصاویر بازیابی‌شده را از دو (چند) روش یادگیری دریافت کرده و با هم ترکیب می‌کند. ترکیب اطلاعات شامل دو بخش است. بخش اول شامل انتخاب روش‌های یادگیری و بخش دوم شامل نحوه ترکیب خروجی روش‌های یادگیری به منظور حصول بهترین نتیجه برای بازیابی تصاویر است. در انتخاب

جدول ۱: مثالی از روش ترکیب کندورست.

داده	رقیب			
	I_1	I_2	I_3	I_4
I_1	-	۰	۰	۱
I_2	۱	-	۰	۱
I_3	۱	۰	-	۱
I_4	۰	۰	۰	-

جدول ۲: مثالی از روش ترکیب کندورست.

داده	رقیب			
	I_1	I_2	I_3	I_4
I_1	-	۲	۱	۲
I_2	۱	-	۱	۲
I_3	۲	۱	-	۲
I_4	۱	۱	۱	-

۳-۲-۱ ترکیب بدون وزن

در این نوع ترکیب که در (۶) نشان داده شده است، می‌توان از روش‌هایی چون حداقل، حداکثر، ضرب، ترکیب خطی، روش شمارش بوردا^۱ (BC)، ترکیب رتبه متقابل^۲ (RRF) و کندورسه^۳ استفاده کرد

$$rank = CombineRanks(Rank_{s_1}, Rank_{s_2}) \quad (۶)$$

- روش شمارش بوردا [۲۵] تا [۲۷] اولین روشی است که از تئوری رأی‌گیری استفاده می‌کند. در این روش اگر در نظر بگیریم که n داده در پایگاه داده ذخیره شده‌اند، به داده‌ای با پایین‌ترین رتبه n رأی اختصاص داده می‌شود و به رتبه‌های بالاتر یک رأی کمتر اختصاص داده می‌شوند (به رتبه دوم $(n-1)$ و به رتبه سوم $(n-2)$ و ...). در آخر تمام رأی‌های به دست آمده از روش‌های یادگیری مختلف برای یک داده با یکدیگر جمع شده و داده‌ای با بالاترین رأی در جایگاه اول قرار می‌گیرد

$$BC(i) = BC_1(i) + BC_2(i) + \dots \quad (۷)$$

$BC_1(i)$ ، $BC_2(i)$ و ... نشان‌دهنده میزان رأی محاسبه‌شده برای داده i ام در روش‌های یادگیری مختلف می‌باشد.

- در روش ترکیب رتبه متقابل (RRF) [۲۸]، داده‌ها با استفاده از رابطه ساده (۸) مرتب‌سازی می‌شوند. در بررسی‌های انجام‌شده $k=60$ انتخاب شده و تغییر نمی‌کند. این رابطه از این واقعیت تبعیت می‌کند که در حالی که داده‌ها با رتبه‌های بالا بسیار مهم هستند، اهمیت داده‌ها با رتبه پایین از بین نمی‌رود. ثابت k تأثیر رتبه‌های بالا را در داده‌های پرت^۴ کم می‌کند

$$RRFscore(i \in I) = \sum_{r \in R} \frac{1}{k+r(i)} \quad (۸)$$

در این رابطه $r(i)$ نشان‌دهنده رتبه تصویر i ام می‌باشد.
 - روش کندورست [۲۹] و [۳۰] از تئوری رأی‌دهی استفاده می‌کند. این الگوریتم یک روش مبتنی بر رأی اکثریت است و نامزدی را به عنوان برنده انتخاب می‌کند که در یک مقایسه از همه نامزدها برود. در این روش، رأی‌دهندگان، نامزدها را بر اساس برتری و اولویت رتبه‌بندی می‌کنند. عملیات رقابت در یک ماتریس انجام می‌گیرد و اطلاعات هر نامزد به صورت یک سطر و ستون در ماتریس قرار می‌گیرد. پس برای m داده یک ماتریس $m \times m$ مورد نیاز است. ابتدا تمام عناصر ماتریس مقدار صفر می‌گیرند. اگر داده i نسبت به داده j برتری داشته باشد (i در رتبه بالاتری نسبت به j قرار

گرفته باشد)، به سطر i و ستون j یک واحد اضافه می‌شود. سپس برای تمام مقادیر a_{ij} ، (۹) بررسی می‌شود. در آخر، امتیاز کلی یک نامزد برابر است با تعداد حالتی که نامزد مربوط در مقایسه جفت نامزدها برنده شده است

$$\begin{cases} \text{if } a_{ij} > \frac{m}{2} \text{ then } d_i^{beats} d_j \\ \text{if } a_{ij} < \frac{m}{2} \text{ then } d_j^{beats} d_i \end{cases} \quad (۹)$$

برای مثال، اگر در نظر بگیریم که ۳ سامانه R_1 ، R_2 و R_3 تصاویر را به صورت $R_1 = \langle I_1, I_2, I_3, I_4 \rangle$ ، $R_2 = \langle I_2, I_3, I_1, I_4 \rangle$ و $R_3 = \langle I_1, I_2, I_3, I_4 \rangle$ مرتب کنند، برای ترکیب آنها از روش کندورست استفاده می‌شود. در سامانه R_1 ، I_2 در رتبه بالاتری نسبت به I_1, I_3, I_4 قرار گرفته و I_1 در رتبه بالاتری نسبت به I_3, I_4 قرار گرفته است و همچنین تصویر I_1 رتبه بالاتری نسبت به تصویر I_2 دارد. به تمام حالت‌های گفته‌شده در ماتریس ۱ واحد اضافه می‌شود و ماتریس به صورت جدول ۱ محاسبه می‌شود.

عملیات را با نتایج به دست آمده از سامانه‌های R_1 و R_2 ادامه می‌دهیم و ماتریس به صورت جدول ۲ به دست می‌آید. در اینجا سه سامانه بازیابی وجود دارد. برای هر سلول $a_{i,j}$ در ماتریس اگر $a_{i,j} \geq 2$ باشد، I_i به I_j برتری دارد. اگر $a_{i,j} \leq 1$ باشد، I_j به I_i برتری دارد. برای تصاویر در هر سطر، تعداد حالت‌هایی که نامزد در مقابل سایر نامزدها برنده شده است، محاسبه می‌گردد و عدد به دست آمده در قسمت امتیاز کلی قرار می‌گیرد. با مرتب‌سازی امتیازهای کلی، تصاویر به صورت $FinalList = \langle I_2, I_1, I_3, I_4 \rangle$ بازیابی می‌شوند.

۳-۲-۲ ترکیب وزن‌دار

در ترکیب اطلاعات می‌توان از رتبه‌بندی وزن‌دار استفاده کرد. در این روش مجموع وزن‌دار رتبه تصاویر مطابق (۱۰) مورد استفاده قرار می‌گیرد

$$FinalRank = (W_1 \times Rank_{s_1}) + (W_2 \times Rank_{s_2}) \quad (۱۰)$$

w نشان‌دهنده وزن مورد استفاده است. در ادامه روش‌های وزن‌دهی مورد استفاده را شرح می‌دهیم.

- استفاده از اطلاعات واریانس، در این حالت طبق (۱۱) از معکوس واریانس رتبه تمام تصاویر مرتبط استفاده می‌شود. هر چه واریانس رتبه تصاویر مرتبط بازیابی‌شده در روش یادگیری کوتاه‌مدت کمتر باشد، وزن اختصاص‌یافته به روش یادگیری بیشتر خواهد بود

$$w = \frac{1}{\text{var}(Rank^+)} \quad (۱۱)$$

- استفاده از روش وزن‌دهی شمیسی [۷] که برای تنظیم وزن‌ها، بازیابی تصویر بر اساس تک‌تک ویژگی‌ها بدون در نظر گرفتن سایر ویژگی‌ها انجام می‌گیرد. سپس رتبه تصاویری که به عنوان مرتبط

1. Borda Count
2. Reciprocal Rank Fusion
3. Condorcet
4. Outlier

w_{ST} نسبت توابع عدم شباهت کوتاه‌مدت را در محاسبه عدم شباهت نهایی کنترل می‌کنند. در [۳۱] و [۳۲] این دو مقدار مساوی در نظر گرفته شده‌اند [۲۴]

$$D(Q, F) = w_{S1} D_{S1}(Q, F) + w_{ST} D_{ST}(Q, F) \quad (13)$$

$$0 \leq w_{S1}, w_{ST} \leq 1, w_{S1} + w_{ST} = 1$$

۴- آزمایش‌ها و نتایج

در این بخش عملکرد سیستم پیشنهادی را ارزیابی کرده و ترکیب روش‌های مختلف را در سطوح متفاوت با هم مقایسه خواهیم کرد.

۴-۱ سامانه پیاده‌سازی شده

سامانه بازیابی تصویر مبتنی بر محتوا در شکل ۱ نمایش داده شده است و جزئیات هر قسمت در ادامه شرح داده خواهد شد.

پایگاه تصویر: در این تحقیق، پایگاه داده ۱۰۰۰۰ تصویری از مجموعه تصاویر کورل مورد استفاده قرار گرفته است. پایگاه فوق شامل ۸۲ گروه معنایی شامل زیر دریا، سرعت، جشنواره، کریسمس، تمبر، دانه‌ها، جوانی، سبزیجات، آبشار، غذا، اسب، چین، تایلند، کوه، فیل، شیر، عقاب، ابر، غروب، پاییز، اهرام، زن، مرد و غیره می‌باشد. از هر گروه معنایی ۱۰۰ یا ۲۰۰ تصویر نمونه با ابعاد 384×256 یا 256×384 در پایگاه وجود دارد که نمونه‌ای از تصاویر پایگاه در شکل ۳ آمده است.

ویژگی‌های دیداری: در این تحقیق در پیاده‌سازی سامانه از سه نوع ویژگی دیداری MPEG-7 [۳۳] استفاده شده است. ویژگی‌های پیاده‌سازی شده عبارت هستند از هیستوگرام رنگ به نمایندگی از ویژگی رنگ، هیستوگرام لبه به نمایندگی از ویژگی شکل و ویژگی گابور به نمایندگی از ویژگی بافت. برای نمایه‌سازی رنگ در تصاویر از هیستوگرام رنگ در فضای HSV استفاده شده است. در این تحقیق هیستوگرام رنگ با ۲۵۶ بازه، ویژگی‌های گابور با ۶۲ ویژگی و ویژگی هیستوگرام لبه با ۸۰ ویژگی از تصاویر استخراج شده و تمام پایگاه تصویر بر این مبنای نمایه‌سازی شده است. در مجموع هر یک از تصاویر با برداری به اندازه ۳۹۸ بعد نمایه‌سازی شده‌اند.

برای نمایه‌سازی رنگ در تصاویر از هیستوگرام رنگ در فضای HSV استفاده شده است. در این تحقیق هیستوگرام رنگ پیشنهادی با ۲۵۶ بازه از تصاویر استخراج شده و تمام پایگاه تصویر بر این مبنای نمایه‌سازی شده است. برای محاسبه هیستوگرام رنگ، هر یک از بازه‌های H ، S و V به ترتیب به ۱۶، ۴ و ۴ بازه به طور خطی و یکنواخت چندی شده است.

لبه، یک ویژگی مهم در بیان محتوای تصویر است. در استاندارد MPEG-7 برای نمایش توزیع لبه در تصویر از هیستوگرام لبه استفاده می‌شود. هیستوگرام توزیع لبه‌های محلی با ۸۰ دسته در MPEG-7 به عنوان توصیف‌گر هیستوگرام لبه (EHD) استاندارد شده که نشان‌دهنده توزیع محلی لبه در تصویر می‌باشد. پس از تقسیم‌بندی تصویر به صورت زیرتصاویر 4×4 ، لبه‌های مربوط به ۴ جهت عمودی، افقی، ۴۵ درجه، ۱۳۵ درجه و یک لبه بدون جهت برای هر زیرتصویر محاسبه می‌شود. با محاسبه هیستوگرام، تعداد $5 \times 16 = 80$ ویژگی به دست می‌آید.

ویژگی‌های گابور به علت خاصیت تفکیک‌پذیری خوب و چندانگانه در حوزه‌های مکان و فرکانس به عنوان ابزاری قوی در تحلیل بافت استفاده می‌شوند. در مجموع از ۳۰ فیلتر گابور در ۵ مقیاس و ۶ جهت استفاده شده است. این فیلترها در ابعاد 256×256 پیکسل طراحی شده‌اند. برای نمایه‌سازی تصویر در مرحله پیش‌پردازش، از تصویر رنگی، تصویر خاکستری ساخته شده و این تصویر به ابعاد 256×256 نرمالیزه می‌شود.



شکل ۳: نمونه‌هایی از تصاویر گروه‌های معنایی در پایگاه تصویر کورل.

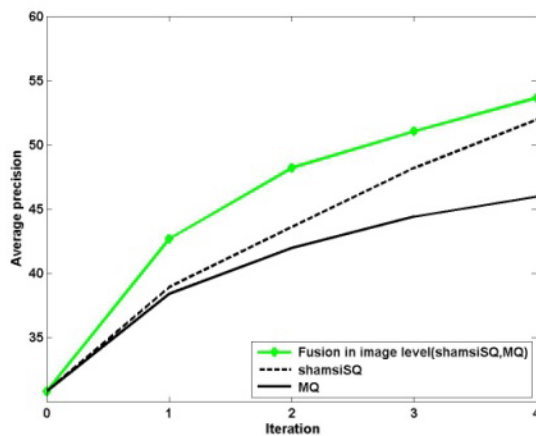
مشخص شده‌اند در بازیابی با هر یک از ویژگی‌ها محاسبه می‌شود. پس از آن وزن مربوط به هر ویژگی طبق (۱۲) متناسب با عکس مجموع رتبه‌های تصاویر مرتبط برای آن ویژگی در نظر گرفته می‌شود. حال برای استفاده از این روش در سامانه پیشنهادی، وزن‌دهی برای هر روش یادگیری در نظر گرفته می‌شود

$$w = \frac{1}{\sum Rank^+} \quad (12)$$

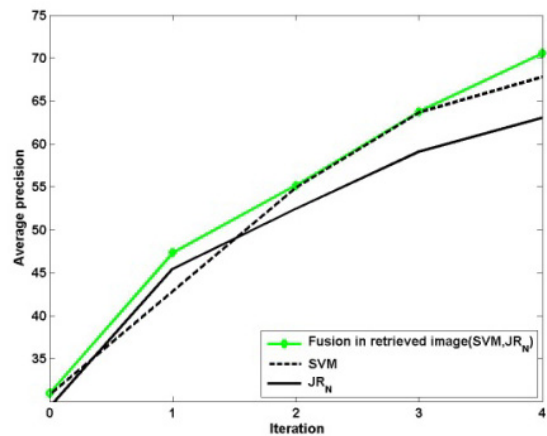
- استفاده از روش وزن‌دهی روی [۱۷]، در این روش سامانه، بازیابی را با استفاده از وزن‌های ثابت و برابر آغاز می‌کند و در پاسخ به تقاضای کاربر، تعداد k تصویر را بر اساس ترکیب ویژگی‌ها، بازیابی کرده و به او ارائه می‌کند. پس از آن که کاربر نظرش را نسبت به تصاویر اعلام کرد، برای تنظیم نقش هر یک از انواع ویژگی، بازیابی تصویر را بر اساس آن نوع ویژگی انجام داده و N تصویر را در لیست آن ویژگی قرار می‌دهد. سپس با مقایسه تصاویر بازیابی شده بر اساس ترکیب ویژگی‌ها و تصاویر بازیابی شده بر اساس آن نوع ویژگی و مشخص کردن تعداد تصاویر مشترک در دو لیست، وزن مربوط به تأثیر نوع ویژگی را با توجه به مرتبط بودن یا نبودن تصاویر مشترک تعیین می‌کند.

۳-۳ ترکیب اطلاعات در سطح تابع شباهت

در این روش، عدم شباهت بین هر دو زوج تصویر به کمک ترکیب عدم شباهت آنها در یادگیری‌های کوتاه‌مدت به دست می‌آید. سپس نتایج به دست آمده مرتب می‌شوند و کمترین عدم شباهت‌ها به عنوان تصاویر بازیابی شده به کاربر نمایش داده می‌شوند. این روش ترکیب اطلاعات در (۱۳) آورده شده است. D_{S1} و D_{ST} نشان‌دهنده توابع عدم شباهت روش‌های یادگیری کوتاه‌مدت هستند. عدم شباهت نهایی که با ترکیب این دو مقدار به دست می‌آید D نام دارد که با کمک یک تابع ترکیب عدم شباهت‌ها به دست می‌آید. جمع وزن‌دار، تابع ترکیبی است که در [۳۱] و [۳۲] برای ترکیب روش‌های یادگیری بلندمدت ارائه شده و برای اولین بار در این مقاله برای ترکیب روش‌های یادگیری کوتاه‌مدت مورد استفاده قرار می‌گیرد. در جمع وزن‌دار، عدم شباهت به کمک ترکیب خطی D_{ST} و D_{S1} مطابق (۱۳) به دست می‌آید. در این روابط وزن‌های w_{S1} و



شکل ۵: ترکیب روش Shamsi-SQ و MQ در سطح تصاویر برگردانده شده.



شکل ۴: ترکیب روش SVM و JR_N در سطح تصاویر برگردانده شده.

روش‌های یادگیری کوتاه‌مدتی که برای ارزیابی ترکیب در سطح تصاویر برگردانده شده در این سامانه مورد استفاده قرار گرفته‌اند عبارتند از طبقه‌بند ماشین بردار پشتیبان [۱۸] و [۱۹]، روش J_{R-N} [۴]، روش تک‌پرسی شمس [۷] و روش چندپرسی [۲۲]. ترکیب روش‌های فوق به صورت زیر می‌باشد:

- ترکیب طبقه‌بند ماشین بردار پشتیبان و روش J_{R-N}

- ترکیب روش تک‌پرسی شمس و روش چندپرسی

شکل ۴ ترکیب روش SVM و J_{R-N} را در سطح تصاویر برگردانده شده نشان می‌دهد که با دقت ۷۰٪، ۳ درصد بهبود نسبت به روش SVM و ۷ درصد بهبود نسبت به روش J_{R-N} داشته است. همچنین شکل ۵ نتایج به دست آمده از ترکیب در سطح تصاویر بازبازی شده را برای روش MQ و شمس به تصویر کشیده است که با دقت ۵۳٪، ۲ درصد بهبود نسبت به روش شمس و ۸ درصد بهبود نسبت به روش MQ داشته است. نتایج به دست آمده، حاکی از برتری روش ترکیب نسبت به روش‌های یادگیری کوتاه‌مدت پایه است.

۴-۲-۲-۴ نتایج ترکیب در سطح رتبه تصاویر

روش‌های یادگیری کوتاه‌مدتی که برای ارزیابی ترکیب در سطح رتبه تصاویر در این سامانه مورد استفاده قرار گرفته‌اند عبارتند از طبقه‌بند ماشین بردار پشتیبان [۱۸] و [۱۹]، روش J_{R-N} [۴]، روش تک‌پرسی شمس [۷] و روش چندپرسی [۲۲]. ترکیب روش‌های فوق به صورت زیر می‌باشد:

- ترکیب طبقه‌بند ماشین بردار پشتیبان و روش J_{R-N}

- ترکیب روش تک‌پرسی شمس و روش چندپرسی

۴-۲-۲-۴-۱ ترکیب بدون وزن در سطح رتبه تصاویر

ترکیب در سطح رتبه بدون وزن در شکل ۶ و برای دو روش SVM و J_{R-N} به تصویر کشیده شده که با استفاده از روش ترکیب RRF و با دقت ۷۳٪، ۶ درصد بهبود نسبت به روش SVM و ۱۰ درصد بهبود نسبت به روش J_{R-N} را نشان می‌دهد.

نتایج به دست آمده در شکل ۷، ترکیب بدون وزن روش شمس و MQ در سطح رتبه تصاویر را نشان می‌دهد که در آن روش ترکیب RRF با دقت ۵۵٪، ۴ درصد بهبود نسبت به روش شمس و ۱۰ درصد بهبود نسبت به روش MQ داشته است.

۴-۲-۲-۴-۲ ترکیب وزن دار در سطح رتبه تصاویر

ترکیب وزن دار در سطح رتبه تصاویر برای دو روش SVM و J_{R-N}

سپس زیرتصویرهای فیلتر شده محاسبه می‌شوند. از هر زیرتصویر فیلتر شده و نیز از خود تصویر، دو پارامتر میانگین و انحراف معیار به عنوان ویژگی استخراج می‌شوند. این ویژگی بر مبنای استاندارد MPEG-7 انتخاب شده است. بردار نمایه ویژگی‌های گابور ۶۲ بعدی است [۴] و [۳۴].

واسط کاربر: برای ارزیابی سامانه و تعامل با کاربر، یک واسط گرافیکی کاربر، پیاده‌سازی شده است که تصویر پرس و جوی کاربر را دریافت کرده و تصاویر مرتبط را برمی‌گرداند. سامانه از طریق واسط طراحی شده، در ۴ تکرار با کاربر تعامل کرده و در هر تکرار، ۲۵ تصویر به وی ارائه می‌کند. کاربر، تصاویر مرتبط و غیر مرتبط را در صفحه گرافیکی علامت‌گذاری و سامانه را مطلع می‌کند.

معیار ارزیابی: معیارهای دقت و فراخوانی از رایج‌ترین معیارهای ارزیابی در بازیابی تصویر هستند. در این سامانه برای ارزیابی روش‌های پیشنهادی از گراف دقت استفاده می‌شود. معیار دقت بیانگر نسبت تصاویر مرتبط بازیابی شده به تصاویر بازیابی شده می‌باشد [۳۵] و [۳۶]

$$\text{precision} = \frac{\text{Number of Relevant Retrieved Images}}{\text{Total Number of Retrieved Images}} \quad (14)$$

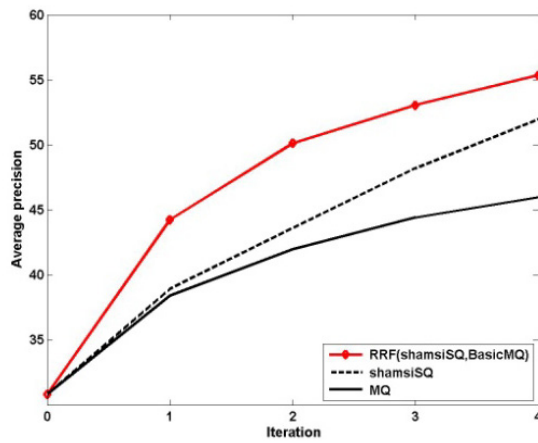
۴-۲-۲-۴-۲ نحوه ارزیابی روش پیشنهادی

برای مقایسه روش‌های یادگیری ۱۰۰ تصویر مختلف انتخاب و به سامانه ارائه شدند. در هر مرحله از بازخورد ربط تعداد تصاویر مرتبط بامعنا کاربر ثبت شده و کاربر با مشخص کردن تصاویر مرتبط و نامرتب بازیابی را ادامه می‌دهد. بر مبنای تعداد تصاویر بازیابی شده مرتبط، برای هر تصویر، گراف دقت بر حسب تعداد تکرارها به دست می‌آید. این عمل برای تمام ۱۰۰ تصویر پرس و جو محاسبه و میانگین آن رسم می‌شود.

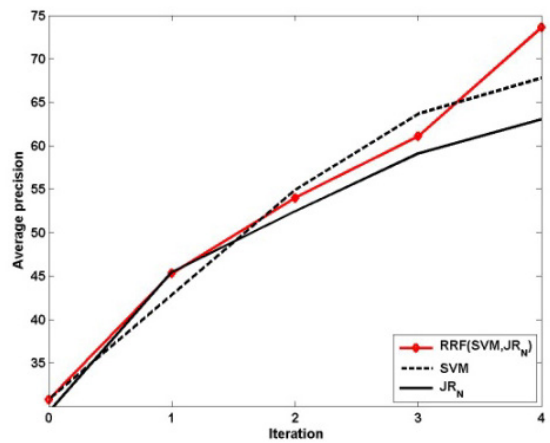
امکان استفاده تمام روش‌های یادگیری کوتاه‌مدت در سه روش ترکیب پیشنهادی وجود ندارد. در حالت‌هایی نیز امکان ترکیب دو روش با یکدیگر ممکن نیست. این موضوع به ذات روش یادگیری کوتاه‌مدت و روش ترکیب برمی‌گردد. این موارد در آزمایش‌ها لحاظ شده است.

۴-۲-۴-۱ نتایج ترکیب در سطح تصاویر برگردانده شده

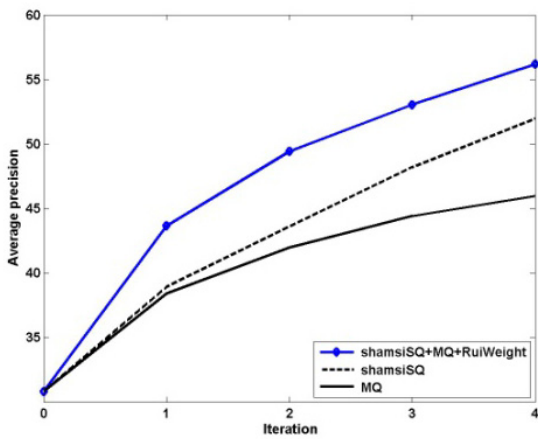
یکی از مراحل مهم در ترکیب اطلاعات، انتخاب مناسب روش‌های یادگیری کوتاه‌مدت پایه می‌باشد. در انتخاب این روش‌ها باید توجه داشت که ترکیب روش‌هایی مناسب است که نتایج یکسان نداشته باشند تا ترکیب نتایج آنها باعث بهبود نرخ بازشناسی سیستم ترکیبی شود. در این راستا می‌توان از روش‌های یادگیری کوتاه‌مدت مبتنی بر طبقه‌بند، روش‌های بهبود بردار پرس و جو، روش‌های بهبود تابع شباهت و روش‌های چندپرسی استفاده کرد.



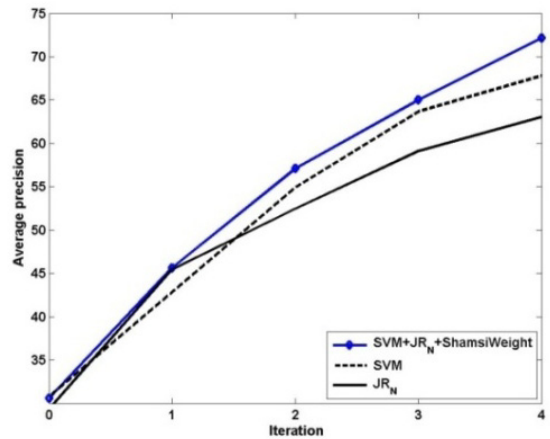
شکل ۷: ترکیب بدون وزن روش Shamsi-SQ و MQ در سطح رتبه تصاویر.



شکل ۸: ترکیب بدون وزن روش SVM و J_{R-N} در سطح رتبه تصاویر.



شکل ۹: ترکیب وزن‌دار روش Shamsi-SQ و MQ در سطح رتبه تصاویر.



شکل ۱۰: ترکیب وزن‌دار روش SVM و J_{R-N} در سطح رتبه تصاویر.

دو روش SVM و J_{R-N} را نمی‌توان در سطح توابع شباهت ترکیب کرد چون فاصله در هر یک معنای متفاوتی دارد. در SVM که روشی مبتنی بر طبقه‌بند می‌باشد، فاصله برای هر تصویر به صورت میزان دور بودن یا نزدیک بودن آن از ابرصفحه محاسبه می‌شود اما در روش J_{R-N} برای به دست آوردن فاصله بین تصاویر پایگاه داده از تصویر پرس و جو، از یک تابع شباهت استفاده می‌شود.

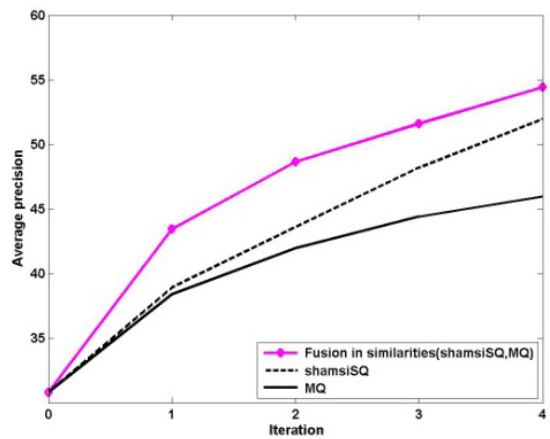
ترکیب دو روش شمس و MQ در سطح تابع شباهت در شکل ۱۰ نمایش داده شده که با دقت ۵۴٪، ۳ درصد بهبود نسبت به روش شمس و ۹ درصد بهبود نسبت به روش MQ داشته است.

۴-۲-۴ تنظیم پارامترها

همان طور که در بخش ۳-۲ شرح داده شد، در ترکیب در سطح رتبه تصاویر (به صورت وزن‌دار و بدون وزن) می‌توان از روش‌های متفاوتی بهره گرفت.

در ترکیب بدون وزن رتبه تصاویر، روش‌هایی چون BC_{max} ، BC_{min} ، ضرب، کندورست و RRF پیاده‌سازی شده و با یکدیگر مقایسه شدند. در ترکیب روش‌های SVM و J_{R-N} (شکل ۱۱)، روش‌های ترکیب RRF و ضرب با دقت ۷۳٪ به روش‌های دیگر برتری داشتند، همچنین روش ترکیب کندورست با دقت ۷۲٪ در جایگاه دوم قرار می‌گیرد و این روش در تکرارهای پایین نیز به خوبی عمل کرده است. در ترکیب روش‌های شمس و چندپرسی (شکل ۱۲)، روش ترکیب RRF با بالاترین دقت به روش‌های دیگر ارجحیت دارد.

به طور کلی در روش ترکیب بدون وزن، روش ترکیب RRF عملکرد بهتری نسبت به روش‌های دیگر از خود نشان داده است. RRF از این



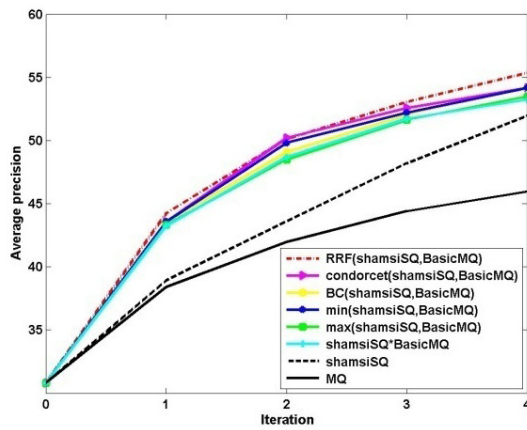
شکل ۱۱: ترکیب روش Shamsi-SQ و MQ در سطح تابع شباهت.

در شکل ۸ با استفاده از روش وزن‌دهی شمس و با دقت ۷۲٪، ۵ درصد بهبود نسبت به روش SVM و ۹ درصد بهبود نسبت به روش J_{R-N} را نشان می‌دهد.

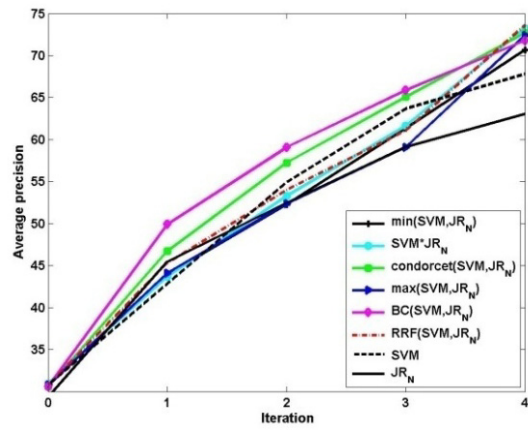
ترکیب وزن‌دار شمس و MQ در سطح رتبه در شکل ۹ با استفاده از روش وزن‌دهی روبی و با دقت ۵۶٪، ۵ درصد بهبود نسبت به روش شمس و ۱۱ درصد بهبود نسبت به روش MQ را نشان می‌دهد.

۴-۲-۳ نتایج ترکیب در سطح توابع شباهت

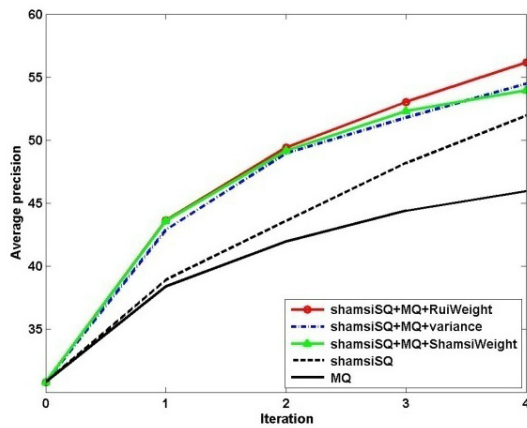
روش‌های یادگیری کوتاه‌مدتی که برای ارزیابی ترکیب در سطح توابع شباهت در این سامانه مورد استفاده قرار گرفته‌اند عبارت هستند از روش تک‌پرسی شمس [۷] و روش چندپرسی [۲۲].



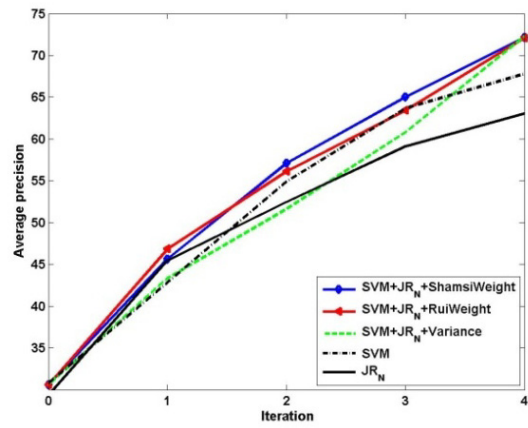
شکل ۱۲: مقایسه روش‌های ترکیب بدون وزن روش Shamsi-SQ و MQ در سطح رتبه تصاویر.



شکل ۱۱: مقایسه روش‌های ترکیب بدون وزن روش SVM و J_{R-N} در سطح رتبه تصاویر.



شکل ۱۴: مقایسه روش‌های ترکیب وزن دار روش Shamsi-SQ و MQ در سطح رتبه تصاویر.



شکل ۱۳: مقایسه روش‌های ترکیب وزن دار روش SVM و J_{R-N} در سطح رتبه تصاویر.

جدول ۴: دقت مربوط به روش‌های یادگیری (شمسی و چندپرسی) و روش‌های ترکیب در هر سه سطح.

روش	دقت
روش شمسی	٪۵۱٫۹۶
روش چندپرسی	٪۴۵٫۹۶
ترکیب بدون وزن در سطح رتبه تصاویر (روش RRF)	٪۵۵٫۳۲
ترکیب وزن دار در سطح رتبه تصاویر (روش وزن دهی رویی)	٪۵۶٫۱۶
ترکیب در سطح تصاویر برگردانده شده	٪۵۳
ترکیب در سطح تابع شباهت	٪۵۴

جدول ۳: دقت مربوط به روش‌های یادگیری (SVM و J_{R-N}) و روش‌های ترکیب در هر سه سطح.

روش	دقت
ماشین بردار پشتیبان	٪۶۷٫۸۰
J_{R-N}	٪۶۳٫۰۴
ترکیب بدون وزن در سطح رتبه تصاویر (روش RRF)	٪۷۳٫۶۰
ترکیب وزن دار در سطح رتبه تصاویر (روش وزن دهی شمسی)	٪۷۲٫۱۲
ترکیب در سطح تصاویر برگردانده شده	٪۷۰

واقعیت تبعیت می‌کند که در حالی که داده‌ها با رتبه‌های بالا بسیار مهم هستند، اهمیت داده‌ها با رتبه پایین از بین نمی‌رود.

در ترکیب وزن دار رتبه تصاویر نیز از روش‌های وزن دهی مختلفی از جمله روش وزن دهی رویی، شمسی و اطلاعات واریانس می‌توان استفاده کرد. این روش‌ها پیاده‌سازی شده و با یکدیگر مقایسه شدند. در ترکیب روش‌های SVM و J_{R-N} (شکل ۱۳)، گرچه در تکرار چهارم تمام روش‌های وزن دهی تقریباً به یک درصد از دقت رسیده‌اند اما روش وزن دهی شمسی در تکرارهای پایین و با وجود داده‌های آموزشی کم نیز به خوبی عمل کرده است. به همین علت، روش وزن دهی شمسی با دقت ٪۷۲ بهترین عملکرد را نشان می‌دهد.

در ترکیب وزن دار روش شمسی و چندپرسی (شکل ۱۴) نیز روش وزن دهی رویی به روش‌های وزن دهی دیگر برتری دارد.

۴-۲-۵ ارزیابی کلی نتایج به دست آمده از روش‌های ترکیب

همان طور که در جدول ۳ نشان داده شده است، روش‌های ترکیب

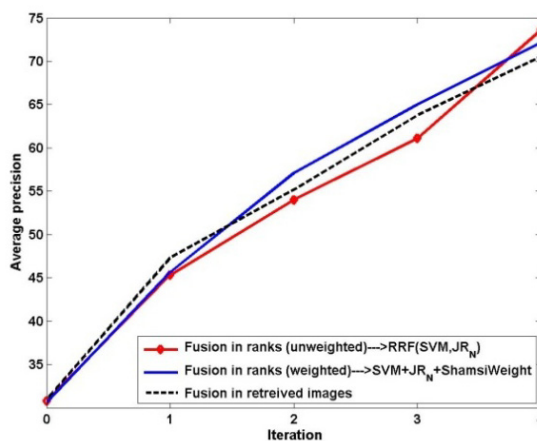
SVM و J_{R-N} در هر سه سطح، بهبود ۳ الی ۶ درصدی نسبت به روش SVM و ۷ الی ۱۰ درصدی نسبت به روش J_{R-N} را در تکرار چهارم داشته‌اند و به طور کلی در ترکیب دو روش ذکر شده، بهترین نتایج مربوط به ترکیب بدون وزن در سطح رتبه تصاویر است که با روش RRF به دست می‌آید و بعد از ۴ تکرار بازخورد ربط، ۶ درصد بهبود نسبت به روش SVM و ۱۰ درصد بهبود نسبت به روش J_{R-N} داشته است (شکل ۱۵). همان طور که در همانطور که در شکل ۱۶ و جدول ۴ مشاهده می‌کنید، نتایج به دست آمده از ترکیب دو روش شمسی و MQ، بهبود ۲ الی ۵ درصدی نسبت به روش شمسی و بهبود ۸ الی ۱۱ درصدی نسبت به روش MQ را در تکرار چهارم داشته‌اند. به طور کلی در این قسمت ترکیب در سطح رتبه وزن دار بهترین نتیجه را دارد که در آن روش وزن دهی RUI با ٪۵۶ دقت، ۵ درصد بهبود نسبت به روش شمسی و ۱۱ درصد بهبود نسبت به روش چندپرسی ارائه می‌کند.

۵- جمع‌بندی

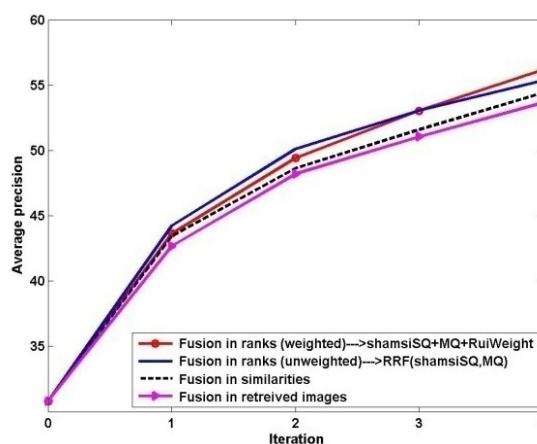
در سامانه‌های بازیابی تصویر بر اساس محتوا از ابزار بازخورد ربط به منظور بالابردن کارایی سامانه و کاهش شکاف معنایی استفاده می‌شود. یادگیری در بازخورد ربط به دو صورت کوتاه‌مدت و بلندمدت انجام می‌گیرد. در این مقاله، رویکرد جدیدی در بازخورد ربط ارائه شده که در آن نتایج به دست آمده از روش‌های یادگیری کوتاه‌مدت در سطوح مختلف و به روش‌های مختلف با یکدیگر ترکیب می‌شوند. واحد ترکیب، چند لیست از اطلاعات را به عنوان ورودی دریافت کرده و تنها یک لیست نهایی از اطلاعات در سه سطح (تصاویر برگردانده‌شده، تابع شباهت و رتبه تصاویر) مورد بررسی قرار گرفته است. روش‌های پیشنهادی در یک پایگاه تصویر با ۱۰۰۰۰ تصویر آزموده و نتایج با روش‌های یادگیری کوتاه‌مدت پایه، مقایسه شده است. در مجموع، ترکیب در سطح تصاویر برگردانده شده، محاسبات چندانی ندارد و نتایج به دست آمده نسبت به روش‌های ترکیب دیگر دقت پایین‌تری دارد (برخوردار برای ویژگی منفی استفاده نمی‌شود). در روش ترکیب در سطح تابع شباهت نیز نتایج قابل قبول است اما به این نکته باید توجه شود که روش‌های یادگیری پایه برای به دست آوردن فاصله تصاویر پایگاه داده از تصویر پرس و جو از تابع شباهت استفاده کنند تا فاصله در هر دو معنای یکسانی داشته باشد. ترکیب در سطح رتبه‌بندی از سایر روش‌های ترکیب برتری دارد و نتایج از دقت بالاتری برخوردار هستند. به طور کلی، نمودارهای به دست آمده نشان می‌دهد که ترکیب روش‌های یادگیری، نتایج بازیابی را نسبت به سامانه با یک روش یادگیری بهبود می‌دهد و گراف دقت در روش‌های پیشنهادی بالاتر از گراف دقت روش‌های یادگیری کوتاه‌مدت قرار گرفته است. نتایج کارآمدبودن روش‌های پیشنهادی را نشان می‌دهد.

مراجع

- [1] R. S. Torres and A. X. Falcao, "Content-based image retrieval: Theory and applications," *RITA*, vol. 13, no. 2, pp. 165-189, 2006.
- [2] S. P. Wilson and G. Stefanou, "Bayesian approaches to content-based image retrieval," in *Proc. of the Int. Workshop/Conf. on Bayesian Statistics and Its Applications*, 11 pp., Varanasi, India, Jan. 2005.
- [3] P. B. Patil and M. B. Kokare, "Relevance feedback in content based image retrieval: A review," *J. of Applied Computer Science & Mathematics*, vol. 5, no. 10, pp. 41-48, 2011.
- [4] E. Rashedi, H. Nezamabadi-pour, and S. Saryazdi, "A Gradient Descent based Similarity Refinement Method for CBIR Systems," Tehran, Iran, 2012.
- [5] X. -Y. Wang, J. -W. Chen, and H. -Y. Yang, "Active SVM-based relevance feedback using multiple classifiers ensemble and features reweighting," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 1, pp. 368-381, Jan. 2012.
- [6] M. E. ElAlami, "Unsupervised image retrieval framework based on rule base system," *Expert Systems with Applications*, vol. 38, no. 4, pp. 3539-3549, Apr. 2011.
- [7] ا. شمسی‌گوشکی، ح. نظام‌آبادی‌پور و س. سریزدی، "بررسی کارایی روش‌های مختلف خوشه‌بندی سلسله‌مراتبی در بازیابی تصویر به شیوه چند پرسشی،" مجموعه مقالات شانزدهمین کنفرانس بین‌المللی سالانه انجمن کامپیوتر ایران، صص. ۳۷۳-۳۶۸، تهران، ۱۹-۱۷ اسفند ۱۳۸۹.
- [8] C. D. Ferreira, et al., "Relevance feedback based on genetic programming for image retrieval," *Pattern Recognition Letters*, vol. 32, no. 1, pp. 27-37, Jan. 2011.
- [9] M. Broilo and F. G. B. De Natale, "A stochastic approach to image retrieval using relevance feedback and particle swarm optimization," *IEEE Trans. on Knowledge and Data Engineering*, vol. 12, no. 4, pp. 267-277, Mar. 2010.
- [10] D. Liu, K. A. Hua, K. Vu, and N. Yu, "Fast query point movement techniques for large CBIR systems," *IEEE Trans. on Knowledge and Data Engineering*, vol. 21, no. 5, pp. 729-743, May 2009.



شکل ۱۵: ارزیابی نتایج به دست آمده از ترکیب SVM و J_{R-N} در هر سه سطح.



شکل ۱۶: ارزیابی نتایج به دست آمده از ترکیب Shamsi-SQ و MQ در هر سه سطح.

اگر بخواهیم روش‌های ترکیب را از لحاظ عملکرد مرتب کنیم، ترکیب در سطح رتبه‌بندی (به صورت وزن‌دار و بدون وزن) در جایگاه اول قرار می‌گیرد. ترکیب در سطح توابع شباهت در جایگاه دوم قرار گرفته و در جایگاه سوم، روش ترکیب در سطح تصاویر بازیابی شده می‌باشد. نتایج نشان می‌دهد که با روش‌های یادگیری ثابت، رویکردهای مختلف ترکیب اطلاعات منجر به نتایج متفاوتی می‌شوند. در مجموع، ترکیب در سطح تابع شباهت محاسبات بیشتری نسبت به سایر روش‌های ترکیب داشته است اما نتایج قابل قبولی دارد. در ترکیب در سطح تصاویر برگردانده شده، تعدادی تصویر به کمک STL_4 و تعدادی به کمک STL_4 برگردانده شدند که محاسبات چندانی ندارد اما نتایج به خوبی روش‌های ترکیب دیگر نیست. ترکیب در سطح رتبه نیز تصاویر را با استفاده از ترکیب رتبه‌بندی به کمک ویژگی‌های سطح پایین به علاوه ویژگی‌های معنایی محاسبه می‌کند که به سایر روش‌های ترکیب ارجحیت دارد و به دقت بالاتری دست پیدا می‌کند.

۴-۲-۶ نتایج پرس و جو

برای مقایسه روش‌های یادگیری کوتاه‌مدت، یک تصویر پرس و جو انتخاب و به سامانه ارائه می‌شود. هر نشست بازیابی تا ۴ مرحله جلو رفته و در هر مرحله تصاویر مرتبط و نامرتب برچسب‌گذاری می‌شوند. در شکل ۱۷ (۳- ب تا ۳- و)، نتایج بازیابی سامانه در بازیابی چهارم با روش‌های SVM و J_{R-N} و ترکیب این دو روش برای تصویر پرس و جو از گروه معنایی سرعت ۳- الف به عنوان نمونه ارائه شده است.



(ب)



(الف)



(د)



(ج)



(و)



(ه)

شکل ۱۷: نتایج بازیابی سامانه در بازیابی چهارم برای تصویر پرس و جوی (الف)، (الف) تصویر پرس و جوی از گروه معنایی سرعت، (ب) نتایج بازیابی با روش SVM، تعداد ۱۸ تصویر مرتبط بازیابی شده است، (ج) نتایج بازیابی با روش J_{R-N} ، تعداد ۹ تصویر مرتبط بازیابی شده است، (د) نتایج بازیابی با ترکیب بدون وزن روش‌های SVM و J_{R-N} در سطح رتبه تصاویر، تعداد ۲۵ تصویر مرتبط بازیابی شده است، (ه) نتایج بازیابی با ترکیب وزن‌دار روش‌های SVM و J_{R-N} در سطح رتبه تصاویر، تعداد ۲۴ تصویر مرتبط بازیابی شده است و (و) نتایج بازیابی با ترکیب روش‌های SVM و J_{R-N} در سطح تصاویر برگردانده‌شده، تعداد ۲۰ تصویر مرتبط بازیابی شده است.

[۱۶] س. ح. نبوی کریزی و ا. کبیر، "یک روش دومرحله‌ای برای ترکیب طبقه‌بندها"، نشریه برق و مهندسی کامپیوتر ایران، سال ۶، شماره ۱، صص. ۶۳-۷۰. بهار ۱۳۸۷.

[17] Y. Rui, T. S. Huang, M. Ortega, and S. Mehrotra, "Relevance feedback: a power tool for interactive content-based image retrieval," *IEEE Trans. Circuits Systems Video Technology*, vol. 8, no. 5, pp. 644-655, Sep. 1998.

[18] S. Tong and E. Chang, "Support vector machine active learning for image retrieval," in *Proc. 9th Int. Conf. ACM Multimedia*, pp. 107-118, 2001.

[19] P. H. Gosselin and M. Cord, "Active learning methods for interactive image retrieval," *IEEE Trans. Image Processing*, vol. 17, no. 7, pp. 1200-1211, Jul. 2008.

[20] J. J. Rocchio, "Relevance feedback in information retrieval," in Salton, G. (Ed.), *The Smart Retrieval System: Experiments in Automatic Document Processing*, Prentice Hall, pp. 313-323, 1971.

[21] Y. Rui, T. S. Huang, and S. Mehrotra, "Content-based image retrieval with relevance feedback in MARS," in *Proc. IEEE Int. Conf. on Image Processing*, vol. 2, pp. 815-818, Santa Barbara, CA, USA, 26-29 Oct. 1997.

[11] M. Arevalillo-Herraez, F. H. Ferri, and S. Moreno-Picot, "Distance-based relevance feedback using a hybrid interactive genetic algorithm for image retrieval," *Applied Soft Computing*, vol. 11, no. 2, pp. 1782-1791, Mar. 2011.

[12] Q. Iqbal and K. Aggarwal, "Feature integration, multi-image queries and relevance feedback in image retrieval," in *Proc. 6th Conf. on Visual Information Systems*, pp. 467-474, Miami, FL, USA, 24-26 Sep. 2003.

[۱۳] ا. شمسی، س. سریزدی، ح. نظام‌آبادی‌پور و م. ش. معین، "روشی جدید در بازخورد ربط برای بازیابی تصویر بر اساس محتوا به شیوه چندپرسشی"، *مجله مهندسی برق دانشگاه تبریز*، جلد ۴۰، شماره ۲، صص. ۶۲-۵۱، پاییز و زمستان ۱۳۸۹.

[14] O. H. Bray, "Information integration for data fusion," *Strategic Business Development*, Sandia National Laboratories, 1997.

[۱۵] س. ح. نبوی کریزی و ا. کبیر، "ترکیب طبقه‌بندها: ایجاد گوناگونی و قواعد ترکیب"، *مجله علوم مهندسی کامپیوتر*، نشریه علمی پژوهشی انجمن کامپیوتر ایران، جلد ۳، شماره ۳، صص. ۹۵-۱۰۷، پاییز ۱۳۸۴.

[۳۴] ع. راشدی و ح. نظام‌آبادی‌پور، "یادگیری بلندمدت مبتنی بر الگوهای معنایی با بهره‌گیری از اطلاعات یادگیری کوتاه‌مدت به روش بهبود تابع شباهت در سامانه‌های بازیابی تصویر"، نشریه مهندسی برق و مهندسی کامپیوتر ایران، سال ۹، شماره ۴، صص. ۲۱۲-۲۰۳، زمستان ۱۳۹۰.

[35] H. Nezamabadipour and E. Kabir, "Concept learning by fuzzy k-NN classification and relevance feedback for efficient image retrieval," *Expert Systems with Applications*, vol. 36, no. 3, pt. 2, pp. 5948-5954, Apr. 2009.

[36] E. Rashedi, H. Nezamabadipour, and S. Saryazdi, "A simultaneous feature adaptation and feature selection method for content-based image retrieval systems," *Knowledge-Based Systems*, vol. 39, pp. 85-94, Feb. 2012.

بهاره باقری مدرک کارشناسی خود را در رشته ی مهندسی کامپیوتر در سال ۱۳۸۹ از دانشگاه علامه جعفری رفسنجان و مدرک کارشناسی ارشد IT-سیستم‌های چند رسانه‌ای را از دانشگاه شهید باهنر کرمان در سال ۱۳۹۲ دریافت نموده است. زمینه‌های تحقیقاتی مورد علاقه‌ی او پردازش تصویر و بازیابی تصویر است.

مریم پورمخ‌آبادی تحصیلات خود را در دوره کارشناسی در دانشگاه شهید باهنر کرمان، در مقطع کارشناسی ارشد در دانشگاه گیلان و در مقطع دکتری در دانشگاه علم و صنعت ایران در رشته مهندسی برق- الکترونیک به‌ترتیب در سال‌های ۱۳۷۵، ۱۳۷۹ و ۱۳۸۹ به پایان رسانده است. وی هم‌اکنون استادیار بخش مهندسی برق دانشگاه شهید باهنر کرمان است. زمینه‌های پژوهشی مورد علاقه‌ی او پردازش سیگنال، پردازش تصویر، الکترونیک نوری و پردازش سیگنال‌های نوری است.

حسین نظام‌آبادی‌پور تحصیلات خود را در دوره کارشناسی مهندسی برق- الکترونیک در دانشگاه شهید باهنر کرمان در سال ۱۳۷۷ و در مقاطع کارشناسی ارشد و دکتری مهندسی برق- الکترونیک از دانشگاه تربیت مدرس به‌ترتیب در سال‌های ۱۳۷۹ و ۱۳۸۳ به پایان رسانده است. وی هم‌اکنون استاد بخش مهندسی برق دانشگاه شهید باهنر کرمان است. زمینه‌های پژوهشی مورد علاقه‌ی او پردازش تصویر، بازشناسی الگو، کاربرد رایانش نرم در پردازش تصویر و روش‌های بهینه سازی ابتکاری است.

[22] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, Academic Press, ISBN:0-21-686140-4. 1999.

[۲۳] ح. نظام‌آبادی‌پور و ا. کبیر، "ارزیابی معیارهای عدم شباهت در بازیابی و طبقه‌بندی تصویر"، *مجله علمی و پژوهشی فنی مدرس*، شماره ۲۲، صص. ۸۹-۹۸، زمستان ۱۳۸۴.

[۲۴] ع. راشدی، بازیابی تصویر مبتنی بر محتوا با استفاده از ترکیب یادگیری کوتاه‌مدت و یادگیری بلندمدت در فضای معناها، رساله دکتری، دانشگاه شهید باهنر کرمان، مرداد ۱۳۹۲.

[25] R. Nuray and F. Can, "Automatic ranking of information retrieval systems using data fusion," *Information Processing and Management*, vol. 42, no. 3, pp. 595-614, 2006.

[26] J. C. de Borda, "Memoire sur les elections au scrutin," in *Histoire de l'Academie Royale des Sciences*, Paris, 1970.

[27] B. Bagheri, M. Pourmahyabadi, and H. Nezamabadipour, "A novel content based image retrieval approach by fusion of short term learning methods," in *Proc. 5th Conf. on Information and Knowledge Technology*, pp. 355-358, Shiraz, Iran, 28-30 May 2013.

[28] G. V. Cormack, C. L. A. Clarke, and S. Butcher, "Reciprocal rank fusion outperforms condorcet and individual rank learning methods," in *Proc. of the 32nd ACM SIGIR*, pp. 758-759, Boston, MA, USA, 19-23 Jul. 2009.

[29] W. Shengli and Z. Xiaoqin, "Condorcet fusion for blog opinion retrieval," in *Proc. 23rd Intl Workshop on Database and Expert Systems Applications DEXA'12*, pp. 156-160, Washington, DC, USA, 3-7 Sep. 2012.

[30] M. Montague and J. Aslam, "Condorcet fusion for improved retrieval," *ACM-CIKM*, 2002.

[31] X. Qi, S. Barrett, and R. Chang, "A noise-resilient collaborative learning approach to content-based image retrieval," *International J. of Intelligent Systems*, vol. 26, no. 12, pp. 1153-1175, Dec. 2011.

[32] J. Han, K. N. Ngan, M. Li, and H. Zhang, "A memory learning framework for effective image retrieval," *IEEE Trans. on Image Processing*, vol. 14, no. 4, pp. 511-524, Apr. 2005.

[33] S. F. Chang, T. Sikora, and A. Puri, "Overview of the MPEG-7 standard," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 11, no. 6, pp. 688-695, Jun. 2001.