

ارزیابی موتورهای جستجو در بازیابی تصاویر بر اساس نمایه‌سازی مبتنی

بر متن و محتوا

یعقوب نوروزی^۱

فرشته سوری^۲

چکیده

هدف: هدف پژوهش حاضر، شناسایی توانمندیها و تفاوت‌های موتورهای جستجوی مورد مطالعه در بازیابی تصاویر بر اساس نمایه‌سازی مبتنی بر متن و محتواست.

روش: پژوهش حاضر از نوع کاربردی است. روش مورد استفاده در این پژوهش از نوع تحقیقات ارزیابانه است. پنج موتور جستجو شامل گوگل، یاهو، پیک‌سرچ، لایکاس و اکسالید که قابل دسترس و سازگار با هدفهای پژوهش بودند، به عنوان جامعه پژوهش انتخاب شدند. جمع‌آوری اطلاعات به شیوه مشاهده و با استفاده از سیاهه محقق ساخته صورت گرفت. سیاهه مورد استفاده شامل ۴۰ مؤلفه بود که سه بخش اصلی (معیارهای ارزیابی موتورهای جستجوی تصاویر، بازیابی تصاویر بر اساس نمایه‌سازی مبتنی بر متن و محتوا) و یازده زیربخش را شامل می‌شود.

یافته‌ها: بر اساس نتایج حاصل، مشخص شد موتور جستجوی گوگل نسبت به دیگر موتورها وضعیت مطلوب‌تری را در بازیابی تصاویر دارد و به عنوان کارآمدترین موتور جستجو در بازیابی تصاویر در این پژوهش شناخته شد. در جایگاه‌های بعدی نیز به ترتیب یاهو، لایکاس، اکسالید و پیک سرچ بر اساس درصد فراوانی کسب شده در معیارهای این پژوهش، رتبه‌بندی شدند. همچنین، با توجه به آزمون فرضیه‌های پژوهش در بازیابی تصاویر بین موتورهای جستجوی مورد مطالعه بر اساس نمایه‌سازی مبتنی بر متن و محتوا، تفاوت معناداری وجود دارد.

بحث و نتیجه‌گیری: بر اساس نتایج پژوهش مشخص شد که موتورهای جستجوی عمومی نسبت به موتورهای جستجوی تخصصی، عملکرد بهتری در بازیابی تصاویر دارند. بنابراین، جستجوگران تصاویر در وب می‌توانند در انتخاب موتور جستجوی متناسب با نیاز خود بهتر تصمیم‌گیری کنند. همچنین، نتایج حاصل می‌تواند به متولیان موتورهای جستجوی مورد مطالعه و طراحان داخلی برای

۱. عضو هیئت علمی گروه علم اطلاعات و دانش‌شناسی دانشگاه قم ynorouzi@gmail.com

۲. دانشجوی کارشناسی ارشد کتابداری و اطلاع‌رسانی دانشگاه آزاد اسلامی واحد همدان و مسئول کتابخانه

مرکزی دانشگاه آزاد اسلامی واحد ملایر librarymalayer@yahoo.com

طراحی بهتر کمک کند. ضمن این که سیاهه تهیه شده در این پژوهش می‌تواند در حوزه‌های مشابه نیز مورد استفاده قرار گیرد.
کلیدواژه‌ها: تصویربایی الکترونیکی، موتورهای جستجوی عمومی، نمایه‌سازی تصاویر، نمایه‌سازی متنی، نمایه‌سازی محتوا.

مقدمه

انسان امروزی در همه جا و در هر شرایطی در محاصره تصاویر قرار دارد و آن را در تمام جنبه‌های زندگی خود به کار می‌گیرد (اسمیت^۱، ۲۰۰۱). چنان که «هو»^۲ (۲۰۰۴) نیز بیان می‌کند، تصاویر ابزار قدرتمندی برای انتقال اطلاعات به شمار می‌روند و در ترکیب با متون می‌توانند اطلاعاتی را که با کلمات قابل توصیف نیستند، به نحو بهتری منتقل کنند. ماهیت چند بُعدی تصاویر باعث می‌شود هر کاربری با توجه به نیاز خود، نگاه ویژه‌ای به یک تصویر داشته باشد، و حتی با هدفی جدا از هدف خالق تصویر، آن را به کارگیرد. این مسئله به‌خصوص زمانی که تصاویر رمزی و نمادین هستند، بسیار جدی‌تر است (کریمی، ۱۳۸۵). «حسن و ژانگ»^۳ (۲۰۰۱) نیز معتقدند در حقیقت، مفاهیم موجود در هر تصویر با توجه به جنبه‌های فرهنگی، آموزشی، پژوهشی، علمی و... متفاوت است. وجود یک تصویر در ترکیب با متن می‌تواند به درک و انتقال مفاهیم موجود در آن کمک کند، حال آن که همان تصویر در متنی دیگر مفهوم دیگری را تداعی می‌کند. از سوی دیگر، با توجه به ارزش ذاتی تصاویر، باید اذعان داشت که نحوه سازماندهی تصاویر تا حدود زیادی میزان بهره‌گیری و نتایج حاصل از آن را مشخص می‌کند. واژگان نمایه‌سازی تصاویر نه تنها باید مواردی مانند عنوان و خالق و ناشر آنها را پوشش دهد، بلکه مفاهیم مورد نظر خالق از خلق آن اثر را نیز باید مدنظر قرار دهد. به عبارت دیگر، وقتی در یک پایگاه اطلاعاتی، یک مقاله کوتاه با حدود ۲۰ واژه نمایه‌سازی می‌شود، تصاویر آن را نیز باید دست کم با همان عمق تجزیه و تحلیل

1. Smith.

2. Hove.

3. Hassan and Zhang.

کرد(رابرتز^۱، ۲۰۰۱). بنابراین، روشن است که نمایه‌سازی تصاویر، فعالیت آسان و ساده نیست. فراهم کردن نقاط دسترسی متعدد برای کاربرانی که قصد دارند تصاویر را بازیابی کنند، مستلزم هزینه فراوان و فعالیت مداوم است (عباس‌پور، ۱۳۸۴). به لحاظ روش، نمایه‌سازی تصاویر دو بُعد اصلی متن و محتوا را دربر می‌گیرد که در هنگام نمایه‌سازی باید به آنها توجه نمود. در نمایه‌سازی مبتنی بر متن، اطلاعات مربوط به متن خود تصویر بر اساس موقعیت‌هایشان رده‌بندی می‌شوند که گروه‌هایی مانند: نام فایل تصویر، عنوان تصویر، موضوع تصویر، برچسب اچ.تی.ام.ال.^۲، سیاه و سفید در برابر رنگی، اندازه تصویر، نوع فایل، اندازه فایل و تاریخ فایل را در بر می‌گیرد (باکستر، ۱۹۹۵). چنان که در نمایه‌سازی مبتنی بر محتوا ویژگی‌های اصلی و اولیه تصویر مانند رنگ، شکل، بافت؛ ویژگی‌های منطقی نظیر موجودیت اشیای نمایش داده شده در تصویر؛ و ویژگی‌های خاص تصویر مانند مفهوم و معنای تصویر را در بر می‌گیرد (چو، ۲۰۰۱).

از سوی دیگر، با ظهور محیط‌های اطلاعاتی جدید، شیوه‌های ذخیره و بازیابی تصاویر از جمله نمایه‌سازی آنها با تغییرات عمده‌ای مواجه شده است. از جمله، ابزارهای قدرتمندی که در این عرصه شروع به فعالیت کرده‌اند، موتورهای جستجو هستند که پس از درک اهمیت تصاویر به عنوان منابع اطلاعاتی ارزشمند، امکان نمایه‌سازی و بازیابی آن را نیز فراهم آوردند و به کاربران این فرصت را دادند تا تصاویر مورد نظر خود را بازیابی و جایابی کنند (حسن و ژانگ، ۲۰۰۱). بنابراین، موتورهای جستجو در حال حاضر نه تنها وسیله‌ای برای دسترسی به اطلاعات متنی هستند، بلکه سازماندهی و بازیابی تصاویر را نیز جزء اهداف خود برای جلب بیشتر رضایت کاربران قرار داده‌اند. اما باید این مسئله را مدنظر قرار داد که توانمندی موتورهای جستجو، به ویژه جامعه مورد مطالعه پژوهش حاضر، در این زمینه چگونه

1. Roberts.
2. HTML.

است و این موتورها تا چه حد توانسته‌اند شیوه‌های مربوط به نمایه‌سازی تصاویر را رعایت کنند. بنابراین، چنان که «کوشا» (۱۳۸۱) بیان می‌دارد، آشنایی با امکانات، قابلیت‌ها و توانایی‌های موتورهای جستجوی اینترنتی اهمیت خاصی دارد. پژوهش حاضر نیز با توجه به اهمیت و ضرورت ارزیابی، موتورهای جستجوی گوگل^۱، یاهو^۲، لایکاس^۳، پیک سرچ^۴ و اکسالد^۵ را در ارزیابی تصاویر بر اساس نمایه‌سازی مبتنی بر متن و محتوا در دستور کار خود قرار داد. البته، در این راستا تاکنون، مطالعات و ارزیابی‌هایی نیز انجام شده است که به طور مستقیم و غیر مستقیم این مسئله را مدنظر قرار داده و بر نقش مهم و ارزنده موتورهای جستجو در ارزیابی تصاویر و شناخت برخی امکانات آنها برای جستجوگران و دیگر کاربران، تأکید کرده‌اند. در این رابطه می‌توان به برخی از این پژوهش‌ها از جمله، مهرآبادی (۱۳۸۵)، منصوری (۱۳۸۷)، اسفندیاری مقدم و بهاری-موفق (۱۳۸۸)، لاکدشتی (۱۳۸۸)، پورسیستانی (۱۳۸۹)، فتحیان (۱۳۹۰)، فرخزاد (۱۳۹۰)، چویی و راسموسن^۶ (۲۰۰۳)، دایامنت^۷ (۲۰۰۷)، گوپین ویرمی و زونلی^۸ (۲۰۰۷)، تیزی^۹ (۲۰۰۸)، اینکپن^{۱۰} و همکاران (۲۰۰۸)، واسیلیوا^{۱۱} (۲۰۰۹)، اوزندی^{۱۲} (۲۰۱۰) و کیدامبی^{۱۳} (۲۰۱۰) اشاره کرد. نیز، در زمینه پژوهش در نمایه‌سازی تصاویر و ارزیابی موتورهای جستجو تاکنون پژوهش‌های مختلفی از زوایای متعدد صورت گرفته

1. <http://www.Google.com>.
2. <http://www.Yahoo.com>.
3. <http://www.Lycos.com>.
4. <http://www.Picsearch.com>.
5. <http://www.Exalead.com>.
6. Choi&Rasmussen.
7. Diamant.
8. Guopin, Jeremy and Xunli.
9. <http://www.tasi.ac.uk/resources/searchengines>.
10. Inkpen& et al.
11. Vassilieva.
12. Ozendi.
13. kidambi.

است. در ادامه، به برخی از این پژوهشها که به لحاظ روش و برخورداری از معیارهای خاص شباهت بیشتری با موضوع پژوهش حاضر دارند، اشاره می‌شود.

در ایران «مهرآبادی» (۱۳۸۵) در پژوهش خود به مقایسه موتورهای کاوش عمومی وب از لحاظ نحوه نمایه‌سازی اطلاعات پرداخت. موتورهای کاوش مورد بررسی در این پژوهش عبارت بودند از: گوگل، آلتاویستا^۱ و آلدوب^۲. بر اساس نتایج حاصل، آلدوب در تمام معیارها بیشترین امتیاز را کسب کرد. «منصوری» (۱۳۸۷) در پژوهش خود تلاش کرد تا ساختاری برای پیاده‌سازی یک سیستم بازیابی تصاویر ارائه دهد. در این سیستم از ویژگیهای رنگ و بافت برای بازیابی تصاویر استفاده و در نهایت مدلی بر این مبنا ارائه شد. «اسفندیاری مقدم و بهاری موفق» (۱۳۸۸) در پژوهشی با هدف ارائه سیاهه واری برای بررسی امکانات جستجو، ۱۹ فراموتور را بررسی و ارزیابی کردند. نتایج به دست آمده از این پژوهش نشان داد تمامی ۱۹ فراموتور بررسی با وجود استفاده از برخی امکانات جستجو، برخی قابلیت جستجوی کلمات فرامتنی در سایتها و نمایش حجم صفحات جستجو شده، در هیچ فراموتوری مشاهده نشد. «لاکدشتی» (۱۳۸۸) در رساله خود با استفاده از روش پیمایشی، مدلی را برای طراحی و پیاده‌سازی روش جدیدی برای بازیابی تصویر، بر اساس محتوا و معانی نهفته در تصویر ارائه کرد. همچنین «پورسیستانی» (۱۳۸۹) پژوهشی را با هدف این که در بازیابی تصاویر باید ابتدا ویژگیهای دیداری آنها استخراج و سپس تصاویر نمایه‌سازی شوند، انجام داد است. در نهایت، یک روش جدید برای استخراج ویژگیهای رنگ و بافت از تصاویر فشرده شده در قلمرو JPEG ارائه داده است. «فتحیان» (۱۳۹۰) نیز پژوهشی را با هدف استفاده از دو ویژگی رنگ و بافت تصاویر با هدف توصیف شباهت تصاویر انجام و ساختاری را برای پیاده‌سازی یک سیستم بازیابی تصاویر بر اساس محتویات بصری با استفاده از یادگیری نظریات کاربران ارائه داد. در نهایت، «فرخزاد» (۱۳۹۰) در پژوهش خود،

1. AltaVista.
2. Alltheweb.

میزان دقت ابرموتور Yippy، در بازیابی تصاویر همراه با متن با استفاده از روش خوشه‌بندی نتایج جستجو را بررسی نمود. از بین خوشه‌های ارائه شده به کاربر، خوشه‌ای با عنوان تصویر وجود دارد که تصاویر همراه با متن را گردآوری می‌کند. تعیین ضریب دقت این ابرموتور با استفاده از فرمول دقت و ارزیابی ربط، انجام گرفت. در خارج از ایران، «چویی و راسموسن» (۲۰۰۳) پرسشهای فرمول‌بندی شده توسط اعضای هیئت علمی و دانش‌آموختگانی که اطلاعات تصویری درباره تاریخ آمریکا را در مجموعه حافظه کتابخانه کنگره آمریکا جستجو می‌کردند، بررسی نمودند. نتایج نشان داد در بیشتر موارد برچسبهای تصاویر با محتوای آنها همخوانی دارد. «دایامنت» (۲۰۰۷) در پژوهش خود، نرم‌افزاری پیوسته برای پردازش اطلاعات دیداری معرفی کرد که با شناسایی و قابلیت حل دشواریهای پردازش ادراکی و شناختی تصاویر، ویژگیهای اساسی آن را توصیف می‌کرد. در پژوهش دیگری که توسط «گوپین، یرمی و زونلی» (۲۰۰۷) انجام گرفت، فناوریهای نوینی معرفی شد که به کاربر اجازه می‌دهد در پایگاه‌های اطلاعاتی تصویرهای بزرگ به مرور و هدایت جست‌وجو پردازد و تصاویری را با تکیه بر روشهای مرور و سپس هدایت سریع و قابل کنترل، بازیابی کند. همچنین «اینکین و همکاران» (۲۰۰۸) در پژوهش خود با بررسی روش خوشه‌بندی خاصی برای بازیابی تصاویر که در قالب خوشه‌بندی سلسله مراتبی ظاهر می‌شود، ابعاد این مسئله را بررسی کردند. نتایج نشان داد روشهای متعدد خوشه‌بندی تصاویر به بازیابی بهتر و دقیق‌تر تصاویر مورد نظر منجر می‌شود. ازسوی دیگر، «واسیلیوا» (۲۰۰۹) در پژوهش خود به نظام بازیابی تصاویر بر مبنای محتوا اشاره می‌کند که به حل بخشی از مشکلات این حوزه نظر دارد و نتیجه پژوهش خود را تحلیل ویژگیهای سطح پایین تصاویر و ساختار خصوصیات برداری، در کنار نمایه‌سازی چندبُعدی، طراحی مناسب رابط کاربر و مصورسازی داده که می‌تواند به بازیابی مؤثر تصاویر منجر شود، بیان نمود. «اوزندی» (۲۰۱۰) با هدف ارائه روشی در بازیابی تصاویر، پژوهشی انجام داد. وی چارچوب تغییر ناپذیری هندسی را ارائه کرد که از مقاطع مخروطی از مرزهای جسم در تصاویر تولید می‌شود و به دلیل خاصیت تغییرناپذیری آن مورد استفاده قرار

گرفته است. وی آن را برای تشخیص و بازیابی تصاویر پیشنهاد داد و در نهایت این روش بازیابی را به نام امضای ثابت تصاویر نامگذاری کرد. «کیدامبی» (۲۰۱۰) در پژوهشی با هدف تعامل انسان و رایانه در نمایه‌سازی تصاویر در موتورهای جستجو، به بررسی سه موتور جستجوی گوگل، یاهو، ام.اس.ان. پرداخت. طبق نتیجه، موتور جستجوی گوگل عملکرد بهتری در بازیابی تصاویر به دست آورد.

از بررسی اجمالی پژوهشهای فوق مشخص می‌شود که بازیابی تصاویر از نظر متخصصان فنی (رایانه و الکترونیک) و ارزیابی موتورهای جستجو از نظر متخصصان علوم انسانی (علم اطلاعات) مورد توجه بوده است، هر چند اغلب پژوهشهای نمایه‌سازی تصاویر از دیدگاه متخصصان رایانه و الکترونیک انجام شده است. نیز ارزیابی موتورهای جستجو بیشتر در زمینه بازیابی اطلاعات متنی است، ولی با ترکیب این دو مقوله فنی و نظری با هم، یعنی، در زمینه بازیابی تصاویر بر اساس نمایه‌سازی آنها در موتورهای جستجو، مجال برای پژوهشهای بیشتر وجود دارد. بنابراین، پژوهش حاضر با هدف شناسایی توانمندیهای موتورهای جستجوی مورد مطالعه در بازیابی تصاویر بر اساس نمایه‌سازی مبتنی بر متن و محتوا اجرا و سعی شد تا به پرسشهای زیر پاسخ داده شود:

۱. قابلیت‌های بازیابی تصاویر در موتورهای جستجوی مورد مطالعه این پژوهش چه تفاوت‌هایی با یکدیگر دارد؟
۲. بازیابی تصاویر در موتورهای جستجوی مورد مطالعه این پژوهش بر اساس نمایه‌سازی مبتنی بر متن، چگونه است؟
۳. بازیابی تصاویر در موتورهای جستجوی مورد مطالعه این پژوهش بر اساس نمایه‌سازی مبتنی بر محتوا چگونه است؟
۴. کارآمدترین موتور جستجو در بازیابی تصاویر بر اساس حوزه‌های مورد مطالعه در این پژوهش کدام است؟

فرضیه‌های پژوهش

۱. در بازیابی تصاویر موتورهای جستجوی مورد مطالعه در این پژوهش بر اساس نمایه‌سازی مبتنی بر متن، تفاوت معناداری وجود دارد.
۲. در بازیابی تصاویر موتورهای جستجوی مورد مطالعه در این پژوهش بر اساس نمایه‌سازی مبتنی بر محتوا، تفاوت معناداری وجود دارد.

روش‌شناسی و جامعه آماری پژوهش

پژوهش حاضر از نوع کاربردی و روش مورد استفاده در آن از نوع تحقیقات ارزیابانه است. بررسی بازیابی تصاویر در موتورهای جستجو به صورت مشاهده مستقیم انجام شد. جامعه آماری پژوهش را موتورهای جستجوی گوگل، یاهو، لایکاس، پیک سرچ و اکسالیید تشکیل می‌دهد، که دو موتور جستجوی عمومی گوگل و یاهو به دلیل جامعیت، و سه موتور جستجوی تخصصی بازیابی چندرسانه‌ای و تصاویر انتخاب شدند. ابزار گردآوری اطلاعات مورد استفاده در این پژوهش، سیاهه محقق ساخته است که بر اساس بررسیهای متون و منابع طراحی شد. همچنین، در تهیه بخش اول این سیاهه یعنی معیارهای ارزیابی موتورهای جستجو، از معیارهای تیزی^۱ (۲۰۰۸) استفاده شد. سیاهه مورد استفاده در مجموع شامل ۴۰ مؤلفه بود که شامل سه بخش اصلی است: معیارهای ارزیابی موتورهای جستجوی تصاویر، بازیابی تصاویر بر اساس نمایه‌سازی مبتنی بر متن و بازیابی تصاویر بر اساس نمایه‌سازی مبتنی بر محتوا نیز در ارزیابی، فقط ۱۰ صفحه اول تصاویر بازیابی شده در هر موتور جستجو بررسی شد. با استفاده از نظرهای متخصصان علم اطلاعات و دانش‌شناسی، از اعتبار (روایی) سیاهه و ارسایی اطمینان حاصل و نظرهای سودمند آنها در سیاهه نهایی اعمال گردید. در واقع، چک لیست متشکل از معیارهاست، بنابراین قابل اعتماد است. همچنین برای اندازه‌گیری پایایی (اعتماد) مقیاس، پس از گردآوری و آماده‌سازی اولیه داده‌های استخراج شده از

1. <http://www.tasi.ac.uk/resources/searchengines>.

_____ ارزیابی موتورهای جستجو در بازیابی تصاویر... / ۱۹۱

سیاهه ارزیابی، برای تجزیه و تحلیل آماری از نرم افزار آماری SPSS^{۱۶} استفاده شد. از جدولهای فراوانی و درصد فراوانی تجمعی برای توصیف سؤالهای سیاهه ارزیابی، و برای آزمون فرضیه‌های تحقیق نیز از آزمون T گروه‌های مستقل استفاده شد.

یافته‌های پژوهش

سؤال اول: قابلیت‌های بازیابی تصاویر در موتورهای جستجوی مورد مطالعه این پژوهش چه تفاوت‌هایی با یکدیگر دارد؟
برای پاسخ به این پرسش، سؤالهای مربوط به معیارهای ارزیابی موتورهای جستجوی تصاویر (دامنه پوشش، گزینه‌های جستجو، عملکرد و کارایی، نحوه نمایش، پشتیبانی و حق مؤلف) در سیاهه ارزیابی بررسی شد که به تفکیک در ادامه می‌آید.

دامنه پوشش

جدول ۱. دامنه پوشش در بازیابی تصاویر در موتورهای جستجوی مورد مطالعه برحسب درصد فراوانی

موتورهای جستجو	بسیار ضعیف	ضعیف	متوسط	خوب	بسیار خوب
ياهو	۵۰	۰	۵۰	۰	۰
پیک سرچ	۵۰	۰	۰	۵۰	۰
گوگل	۰	۰	۰	۰	۱۰۰
لایکاس	۱۰۰	۰	۰	۰	۰
اکسالید	۱۰۰	۰	۰	۰	۰

چنان که از جدول ۱ و درصد فراوانی دامنه پوشش در هر یک از موتورهای جستجو مورد مطالعه پژوهش مشهود، دامنه پوشش در گوگل با ۱۰۰٪ فراوانی بسیار خوب و در لایکاس و اکسالید با ۱۰۰٪ فراوانی بسیار ضعیف است.

گزینه‌های جستجو

جدول ۲. گزینه‌های جستجو در بازیابی تصاویر در موتورهای جستجوی مورد مطالعه برحسب درصد فراوانی

موتورهای جستجو	بسیار ضعیف	ضعیف	متوسط	خوب	بسیار خوب
ياهو	۰	۰	۳۳	۰	۶۷
پیک سرچ	۰	۳۳	۳۳	۳۳	۰
گوگل	۰	۰	۰	۰	۱۰۰
لایکاس	۳۳	۳۳	۳۳	۰	۰
اکسالید	۰	۰	۰	۰	۱۰۰

چنان که در جدول ۲ و وضعیت گزینه‌های جستجو در هر یک از موتورهای جستجوی مورد مطالعه مشهود است، گزینه‌های جستجو در گوگل و اکسالید با ۱۰۰٪ فراوانی بسیار خوب و در لایکاس با ۳۳٪ فراوانی بسیار ضعیف است.

عملکرد و کارایی

جدول ۳. عملکرد و کارایی در بازیابی تصاویر در موتورهای جستجوی مورد مطالعه برحسب درصد فراوانی

موتورهای جستجو	بسیار ضعیف	ضعیف	متوسط	خوب	بسیار خوب
ياهو	۱۲.۵	۰	۱۲.۵	۲۵	۵۰
پیک سرچ	۱۲.۵	۵۰	۳۷.۵	۰	۰
گوگل	۰	۰	۱۲.۵	۲۵	۶۲.۵
لایکاس	۰	۱۲.۵	۵۰	۲۵	۱۲.۵
اکسالید	۰	۵۰	۲۵	۱۲.۵	۱۲.۵

چنان که در جدول ۳ و درصد فراوانی عملکرد و کارایی در هر یک از موتورهای جستجو مورد مطالعه مشاهده می‌شود، عملکرد و کارایی در گوگل با ۶۲.۵٪ فراوانی بسیار خوب، ولی در یاهو و پیک سرچ با ۱۲.۵٪ فراوانی بسیار ضعیف است.

نحوه نمایش

جدول ۴. نحوه نمایش در بازیابی تصاویر در موتورهای جستجوی مورد مطالعه برحسب درصد فراوانی

موتورهای جستجو	بسیار ضعیف	ضعیف	متوسط	خوب	بسیار خوب
ياهو	۵۰	۰	۰	۰	۵۰
پیک سرچ	۵۰	۰	۰	۳۳	۱۷
گوگل	۳۳	۰	۰	۰	۶۷
لایکاس	۱۷	۱۶	۰	۱۷	۵۰
اکسالیید	۵۰	۰	۰	۰	۵۰

جدول ۴ و درصد فراوانی نحوه نمایش در بازیابی تصاویر در موتورهای جستجوی مورد مطالعه پژوهش نشان می‌دهد که نحوه نمایش در گوگل با ۶۷٪ فراوانی بسیار خوب و در یاهو و پیک سرچ و اکسالیید با ۵۰٪ فراوانی بسیار ضعیف و خوب است.

پشتیبانی

جدول ۵. پشتیبانی در بازیابی تصاویر در موتورهای جستجوی مورد مطالعه برحسب درصد فراوانی

موتورهای جستجو	بسیار ضعیف	ضعیف	متوسط	خوب	بسیار خوب
ياهو	۵۰	۵۰	۰	۰	۰
پیک سرچ	۵۰	۵۰	۰	۰	۰
گوگل	۰	۵۰	۰	۵۰	۰
لایکاس	۵۰	۵۰	۰	۰	۰
اکسالیید	۵۰	۵۰	۰	۰	۰

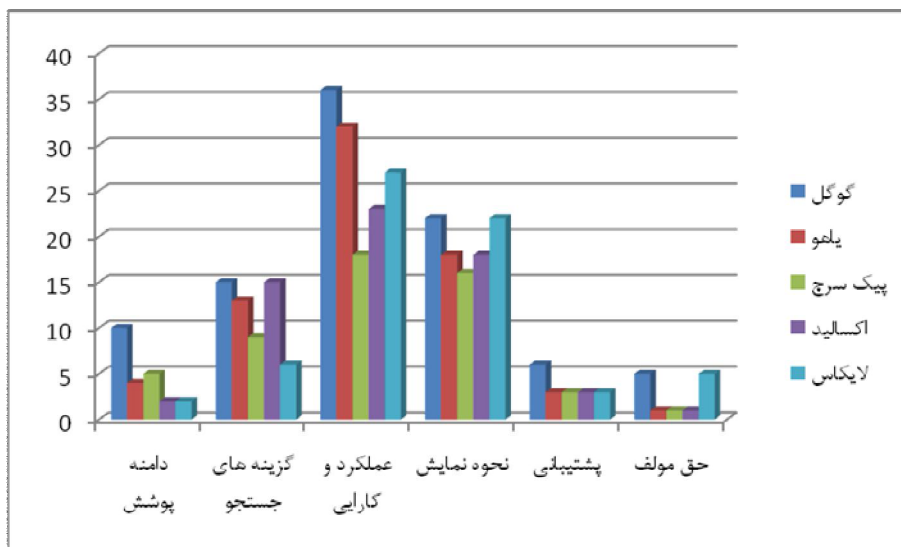
براساس داده های جدول ۵، پشتیبانی در بازیابی تصاویر در گوگل با ۵۰٪ فراوانی خوب و در بقیه موتورها با ۵۰٪ فراوانی، بسیار ضعیف و ضعیف است.

حق مؤلف

جدول ۶. حق مؤلف در بازیابی تصاویر در موتورهای جستجوی مورد مطالعه برحسب درصد فراوانی

موتورهای جستجو	بسیار ضعیف	ضعیف	متوسط	خوب	بسیار خوب
ياهو	۱۰۰	۰	۰	۰	۰
پیک سرچ	۱۰۰	۰	۰	۰	۰
گوگل	۰	۰	۰	۰	۱۰۰
لایکاس	۰	۰	۰	۰	۱۰۰
اکسالید	۱۰۰	۰	۰	۰	۰

از نتایج جدول ۶ و درصد فراوانی حق مؤلف در هر یک از موتورهای جستجوی مورد مطالعه پژوهش مشهود است که رعایت حق مؤلف در گوگل و لایکاس با ۱۰۰٪ فراوانی بسیارخوب و در یاهو و پیک سرچ و اکسالید با ۱۰۰٪ فراوانی، بسیار ضعیف است. در ادامه، در نمودار ۱ تفاوت بین موتورهای جستجو در قابلیت‌های بازیابی تصاویر و این که هر یک از موتورهای جستجو دارای چه وضعیتی هستند، ارائه شده است.



نمودار ۱. وضعیت موتورهای جستجوی مورد مطالعه در قابلیت‌های بازیابی تصاویر

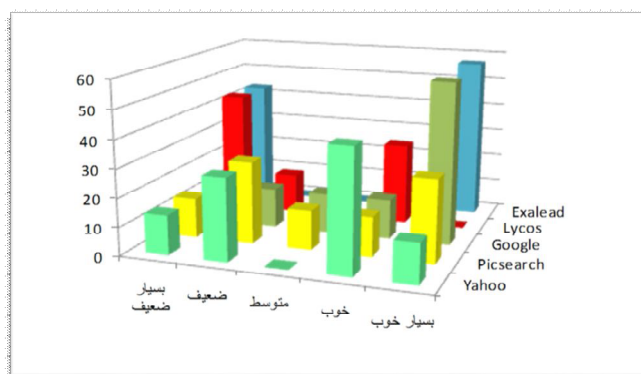
_____ ارزیابی موتورهای جستجو در بازیابی تصاویر... / ۱۹۵

سؤال دوم: بازیابی تصاویر در موتورهای جستجوی مورد مطالعه این پژوهش بر اساس نمایه‌سازی مبتنی بر متن چگونه است؟
برای پاسخ به پرسش، مؤلفه‌های مربوط به نمایه‌سازی مبتنی بر متن تجزیه و تحلیل می‌شود.

جدول ۷. اطلاعات متنی در بازیابی تصاویر در موتورهای جستجوی مورد مطالعه برحسب درصد فراوانی

موتورهای جستجو	بسیار ضعیف	ضعیف	متوسط	خوب	بسیار خوب
ياهو	۱۴	۲۹	۰	۴۳	۱۴
پیک سرچ	۱۴	۲۹	۱۴	۱۴	۲۹
گوگل	۰	۱۴	۱۴	۱۴	۵۷
لایکاس	۴۳	۱۴	۱۴	۲۹	۰
اکسالیید	۴۳	۰	۰	۰	۵۷

چنان که از جدول ۷ و درصد فراوانی نمایه‌سازی مبتنی بر متن در هر یک از موتورهای جستجوی مورد مطالعه پژوهش مشهود است، اطلاعات متنی تصاویر در گوگل و اکسالیید با ۵۷٪ فراوانی بسیار خوب ولی در لایکاس با ۴۳٪ فراوانی، بسیار ضعیف است. همچنین، مقایسه معیارهای بازیابی تصاویر بر اساس نمایه‌سازی مبتنی بر متن در موتورهای جستجوی مورد مطالعه، در نمودار ۲ نیز قابل مشاهده است.



نمودار ۲. مقایسه معیارهای بازیابی تصاویر بر اساس نمایه‌سازی مبتنی بر متن در موتورهای جستجوی مورد مطالعه

سؤال سوم: بازیابی تصاویر در موتورهای جستجوی مورد مطالعه این پژوهش بر اساس نمایه‌سازی مبتنی بر محتوا، چگونه است؟
در این قسمت مؤلفه‌های مربوط به موضوع و مفهوم، رنگ، شکل و بافت تصاویر در موتورهای جستجو تجزیه و تحلیل می‌شود.

موضوع و مفهوم

جدول ۸. موضوع و مفهوم در بازیابی تصاویر در موتورهای جستجوی مورد مطالعه برحسب

درصد فراوانی

موتورهای جستجو	بسیار ضعیف	ضعیف	متوسط	خوب	بسیار خوب
ياهو	۰	۰	۰	۰	۱۰۰
پیک سرچ	۰	۳۳	۶۷	۰	۰
گوگل	۰	۰	۰	۰	۱۰۰
لایکاس	۰	۰	۰	۳۳	۶۷
اکسالیید	۰	۰	۳۳	۳۳	۳۳

چنان که در جدول ۸ مشاهده می‌شود، وضعیت موضوع و مفهوم در گوگل و یاهو با ۱۰۰٪ فراوانی بسیار خوب، ولی در پیک سرچ با ۳۳٪ ضعیف است.

رنگ

جدول ۹. رنگ در بازیابی تصاویر در موتورهای جستجوی مورد مطالعه برحسب درصد فراوانی

موتورهای جستجو	بسیار ضعیف	ضعیف	متوسط	خوب	بسیار خوب
ياهو	۰	۳۳	۰	۶۷	۰
پیک سرچ	۳۳	۰	۳۳	۰	۳۳
گوگل	۰	۰	۳۳	۳۳	۳۳
لایکاس	۰	۰	۱۰۰	۰	۰
اکسالیید	۳۳	۰	۳۳	۰	۳۳

_____ ارزیابی موتورهای جستجو در بازیابی تصاویر... / ۱۹۷

داده‌های جدول ۹ نشان می‌دهد وضعیت رعایت رنگ در گوگل با ۳۳٪ در حد متوسط و یاهو با ۶۷٪ خوب است. همچنین، در اکسالیید و پیک سرچ با ۳۳٪ فراوانی در حد متوسط است.

شکل

جدول ۱۰. شکل در بازیابی تصاویر در موتورهای جستجوی مورد مطالعه برحسب درصد فراوانی

موتورهای جستجو	بسیار ضعیف	ضعیف	متوسط	خوب	بسیار خوب
ياهو	۳۳	۳۳	۳۳	۰	۰
پیک سرچ	۶۷	۳۳	۰	۰	۰
گوگل	۳۳	۰	۳۳	۳۳	۰
لایکاس	۳۳	۰	۳۳	۳۳	۰
اکسالیید	۶۷	۰	۳۳	۰	۰

طبق جدول ۱۰، رعایت معیارهای شکل در گوگل و لایکاس در مقایسه با بقیه خوب و در پیک سرچ و اکسالیید بسیار ضعیف است.

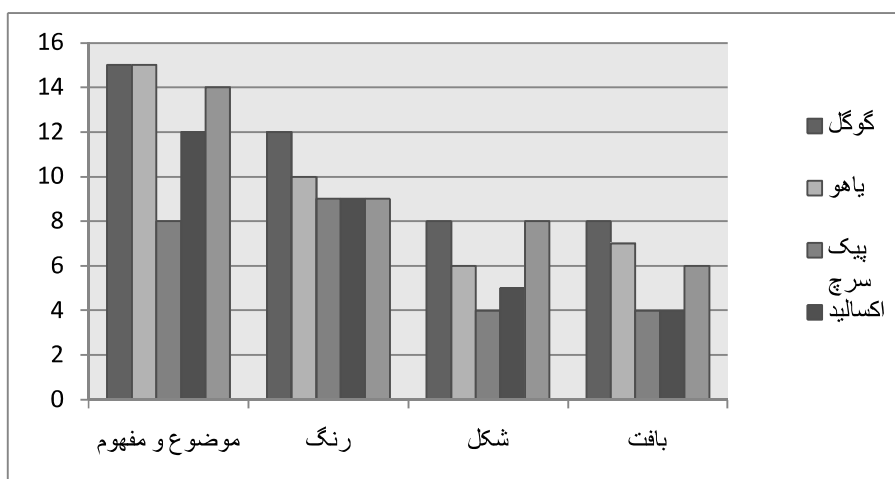
بافت

جدول ۱۱. بافت در بازیابی تصاویر در موتورهای جستجوی مورد مطالعه برحسب درصد فراوانی

موتورهای جستجو	بسیار ضعیف	ضعیف	متوسط	خوب	بسیار خوب
ياهو	۰	۰	۵۰	۵۰	۰
پیک سرچ	۵۰	۰	۵۰	۰	۰
گوگل	۰	۰	۵۰	۰	۵۰
لایکاس	۰	۰	۱۰۰	۰	۰
اکسالیید	۵۰	۰	۵۰	۰	۰

چنان که از جدول ۱۱ مشهود است، رعایت معیارهای بافت در گوگل با ۵۰٪ فراوانی به سمت بسیار خوب تمایل دارد، و در پیک سرچ و اکسالیید با ۵۰٪ به سمت بسیار ضعیف و متوسط است.

در نتیجه، طبق تجزیه و تحلیل مؤلفه‌های بازیابی تصاویر بر اساس نمایه‌سازی مبتنی بر محتوا (۱۱ مؤلفه) در هر یک از موتورهای جستجوی مورد مطالعه پژوهش به صورت جداگانه، دریافت گردید که موتور جستجوی گوگل با ۴۵٪ به سمت بسیار خوب و یاهو با ۵۲٪ به سمت خوب و لایکاس با ۵۵٪ به سمت متوسط است. پیک‌سرچ و اکسالید با ۷۲٪ به سمت بسیار ضعیف می‌باشند. نیز، در نمودار ۳ وضعیت بازیابی تصاویر در موتورهای جستجوی مورد مطالعه این پژوهش بر اساس نمایه‌سازی مبتنی بر محتوا، بر اساس محاسبه میانگین نمایان است.



نمودار ۳. وضعیت بازیابی تصاویر در موتورهای جستجوی مورد مطالعه این پژوهش بر اساس نمایه‌سازی مبتنی بر محتوا بر اساس میانگین

سؤال چهارم: کارآمدترین موتور جستجو در بازیابی تصاویر بر اساس حوزه‌های

مورد مطالعه در این پژوهش کدام است؟

به منظور پاسخ به پرسش چهارم، کل مؤلفه‌های بازیابی (۴۰ مؤلفه) برای هر یک از موتورهای جستجو مورد مطالعه پژوهش با استفاده آزمون کروسکال والیس به صورت جداگانه تجزیه و تحلیل شد. نتایج در جدول ۱۳ ارائه شده است.

_____ ارزیابی موتورهای جستجو در بازیابی تصاویر... / ۱۹۹

جدول ۱۳. رتبه میانگین موتورهای جستجوی در بازیابی تصاویر بر اساس حوزه‌های مورد مطالعه

رتبه میانگین	تعداد سؤالات	موتورهای جستجو
۱۰۵.۱۶	۴۰	ياهو
۷۶.۲۴	۴۰	پیک سرچ
۱۳۶.۶۲	۴۰	گوگل
۹۲.۸۸	۴۰	لایکاس
۹۱.۶۰	۴۰	اکسالید

جدول ۱۴. تست آماره آمون برای کارآمدترین موتور جستجو در بازیابی تصاویر بر اساس

حوزه‌های مورد مطالعه

تست آماره آمون	
۲۵.۸۸۵	کای اسکور
۴	درجه آزادی
۰.۰۰	سطح معنادار

بر اساس جدول ۱۳ نتیجه‌گیری می‌شود که موتور جستجوی گوگل بیشترین رتبه میانگین را به خود اختصاص داده است و پس از گوگل به ترتیب یاهو، لایکاس، اکسالید و پیک سرچ قرار دارند. همچنین، با توجه به سطح معناداری به دست آمده در جدول ۱۴ (۰.۰۰)، چون آلفای مفروض از ۰.۰۵ کوچکتر است، پس با اطمینان ۰.۹۵ درصد نتیجه‌گیری می‌شود که کارآمدترین موتور جستجو، گوگل است.

پاسخ به فرضیه‌های پژوهش

فرضیه اول: در بازیابی تصاویر موتورهای جستجوی مورد مطالعه این پژوهش بر اساس نمایه‌سازی مبتنی بر متن، تفاوت معناداری وجود دارد.

H_0 : در بازیابی تصاویر موتورهای جستجوی مورد مطالعه این پژوهش بر اساس

نمایه‌سازی مبتنی بر متن تفاوت معناداری وجود ندارد.

H_1 : در بازیابی تصاویر موتورهای جستجوی مورد مطالعه این پژوهش بر اساس نمایه‌سازی مبتنی بر متن تفاوت معناداری وجود دارد.

جدول ۱۳. بررسی اختلاف بین بازیابی موتورهای جستجو بر اساس نمایه‌سازی مبتنی بر متن با

استفاده از آزمون تی (T) گروه‌های مستقل

متغیر	میانگین	واریانس	آماره تی	درجه آزادی	سطح معنادار
موتورهای جستجو	۶۸.۸۰	۱۵.۶۷	۶.۳۵	۸	۰.۰۰۰
نمایه‌سازی مبتنی بر متن	۲۲.۴۰	۴.۶۱			

چنان که در جدول ۱۳ ملاحظه می‌شود، چون سطح معناداری به دست آمده ۰.۰۰ از آلفای مفروض ۰.۰۵ کوچکتر است، پس فرض صفر با اطمینان ۰.۹۵ رد و فرض یک تأیید می‌شود؛ یعنی در بازیابی تصاویر بین موتورهای جستجوی مورد مطالعه این پژوهش بر اساس نمایه‌سازی مبتنی بر متن، تفاوت معناداری وجود دارد.

فرضیه دوم: در بازیابی تصاویر موتورهای جستجوی مورد مطالعه این پژوهش بر اساس نمایه‌سازی مبتنی بر محتوا، تفاوت معناداری وجود دارد.

H_0 : در بازیابی تصاویر موتورهای جستجوی مورد مطالعه این پژوهش بر اساس نمایه‌سازی مبتنی بر محتوا تفاوت معناداری وجود ندارد.

H_1 : در بازیابی تصاویر موتورهای جستجوی مورد مطالعه این پژوهش بر اساس نمایه‌سازی مبتنی بر محتوا تفاوت معناداری وجود دارد.

جدول ۱۴. بررسی اختلاف بین بازیابی موتورهای جستجو بر اساس نمایه‌سازی مبتنی بر محتوا با

استفاده از آزمون تی (T) گروه‌های مستقل

متغیر	میانگین	واریانس	آماره تی	درجه آزادی	سطح معنادار
موتورهای جستجو	۶۸.۸۰	۱۵.۶۷	۴.۵۰	۸	۰.۰۰۲
نمایه‌سازی مبتنی بر محتوا	۳۴.۴۰	۶.۸۰			

چنان که در جدول ۱۴ ملاحظه می‌شود، چون سطح معناداری به دست آمده ۰.۰۰۲ از آلفای مفروض ۰.۰۵ کوچکتر است، پس فرض صفر با اطمینان ۰.۹۵ رد و فرض یک

تأیید می‌شود؛ یعنی در بازیابی تصاویر بین موتورهای جستجوی مورد مطالعه در این پژوهش بر اساس نمایه‌سازی مبتنی بر محتوا تفاوت معناداری وجود دارد.

بحث و نتیجه‌گیری

با توجه به اهمیت بازیابی اطلاعات در دنیای وب و ظهور حوزه‌های جدید بازیابی اطلاعات از جمله تصاویر، در پژوهش حاضر وضعیت موتورهای جستجوی مورد مطالعه در بازیابی تصاویر بر اساس نمایه‌سازی مبتنی بر متن و محتوا ارزیابی شد. بر اساس یافته‌های پژوهش، مؤلفه‌های بازیابی تصاویر در موتورهای جستجو را می‌توان به شرح زیر، مورد بحث و بررسی قرار داد. در بررسی و ارزیابی مؤلفه‌های فرعی ملاحظه شد که توجه کمی به ذکر تاریخ روزآمدسازی شده است. اما در مقابل با توجه به این نکته که اولویت موتورهای جستجو در بازیابی اطلاعات، برخوردار بودن از امکانات و منوهای جستجوی مناسب است و نیز بر اساس این که جستجوی تصاویر و اطلاعات متنی هر دو از یک درگاه انجام می‌گیرد، نتایج حاصل وضعیت گزینه‌های جستجو را در موتورهای جستجو بسیار خوب نشان داده است که با یافته‌های پژوهش «اسفندیاری مقدم و بهاری موفق» (۱۳۸۸) نیز همسو است. همچنین در مؤلفه‌های ربط تصاویر بازیابی شده با جستجوی کاربر، و اعمال پیوند در موتورهای جستجو وضعیت ضعیفی مشاهده شد. این نکته نیز قابل توجه است که اگر پشتیبانی اطلاعات در موتورهای جستجوی تصاویر بیشتر مورد توجه قرار گیرد، در حوزه علوم به خصوص علمی مانند پزشکی، معماری، هنر و تاریخ بسیار کاربرد دارد، زیرا در این حوزه‌ها بازیابی تصویر مناسب می‌تواند نقش بسزایی در کمک به کاربران داشته باشد. همچنین نتایج نشان داد رعایت حق مؤلف تصاویر در موتورهای جستجوی تخصصی مورد مطالعه، برخلاف موتورهای عمومی وضعیت مناسبی ندارد که این مسئله جای تأمل بسیاری دارد و نیازمند پژوهش‌های بیشتری است. اما در کل، نتایج نشان داد از بین جامعه آماری مورد مطالعه، گوگل با کسب رتبه میانگین ۱۳۶.۶۲ و یاهو با ۱۰۵.۱۶ و لایکاس با ۹۲.۸۸ و اکسالید با ۹۱.۶۰ در نهایت پیک سرچ با ۷۶.۲۴ در رتبه اول تا

پنجم قرار دارند. همچنین، در بررسی مؤلفه‌های بازیابی تصاویر بر اساس نمایه‌سازی مبتنی بر متن، نتایج به دست آمده نشان داد در موتورهای جستجوی مورد مطالعه سعی شده است تمامی اطلاعات بصری و متنی تصاویر نمایه‌سازی شود. بنابراین، یافته‌های پژوهش حاضر با پژوهش «جویی و راسموسن» (۲۰۰۳)، پژوهش «فرخ‌زاد» (۱۳۹۰) و «اینکپن و همکاران»^۱ (۲۰۰۸) همسوست، که نشان از وضعیت بسیار خوب بازیابی تصاویر بر اساس نمایه‌سازی مبتنی بر متن است. ضمن این که نتایج نشان داد موتور جستجوی گوگل و اکسالیید با ۵۷٪ به سمت بسیار خوب، یاهو با ۴۳٪ به سمت خوب و پیک سرچ با ۵۸٪ به سمت متوسط و لایکاس با ۴۳٪ به سمت بسیار ضعیف است. همچنین، نتایج یافته‌های پژوهش در مؤلفه‌های بازیابی تصاویر بر اساس نمایه‌سازی مبتنی بر محتوا نشان داد که وضعیت موتور جستجوی گوگل با ۴۵٪ به سمت بسیار خوب و یاهو با ۵۲٪ به سمت خوب و لایکاس با ۵۵٪ به سمت متوسط است. نیز در پیک سرچ و اکسالیید با ۷۲٪ به سمت بسیار ضعیف است. ضمن این که در این زمینه در مؤلفه موضوع و مفهوم، نتایج پژوهش حاضر با یافته‌های پژوهش «لاکدشتی» (۱۳۸۸) همسوست. همچنین، با توجه به اهمیت رنگ در تصاویر به عنوان یک مقوله کاملاً بصری، متأسفانه بازیابی تصاویر بر اساس رنگ تصاویر در موتورهای جستجو وضعیت متوسطی دارد و ضروری است موتورهای جستجو بیشتر از قبل به این مؤلفه از محتوای تصاویر توجه کنند. البته، امکان ترسیم خود شکل در موتورهای جستجو برای بازیابی تصاویر به جای کلیدواژه ضروری است، زیرا برخی از اشکال زاویه‌دار و پیچیده (مانند شکل گلبرگ گلها) از طریق کلیدواژه قابل جستجو نیستند. همچنین، امکان ترسیم شکل برای کاربرانی مانند معماری، گیاه‌شناسی بسیار کاربردی و سودمند است، هر چند در این زمینه نیز نتایج به سمت متوسط و بسیار ضعیف تمایل دارد. ذکر این نکته حایز اهمیت است که بیشتر پژوهش‌های صورت گرفته در نمایه‌سازی محتوایی تصاویر، توسط متخصصان رایانه و الکترونیک صورت گرفته که این مسئله نشانگر

_____ ارزیابی موتورهای جستجو در بازیابی تصاویر... / ۲۰۳

اهمیت بالای این حوزه است. ضمن این که نتایج برخی از این پژوهشها مانند «منصوری» (۱۳۸۷)، «فتحیان» (۱۳۹۰) و «اوزندی»^۱ (۲۰۱۰) با نتایج پژوهش حاضر همسوست.

در کل، نتایج تحلیلهای آماری داده‌های پژوهش بیانگر این است که در بازیابی تصاویر موتورهای جستجوی مورد مطالعه این پژوهش بر اساس نمایه‌سازی مبتنی بر متن و محتوا تفاوت معناداری وجود دارد. همچنین، موتور جستجوی گوگل در تمام معیارهای مورد پژوهش بیشترین امتیاز را کسب کرد و کارآمدترین موتور جستجو در این پژوهش شناخته شد که این مسئله با یافته‌های پژوهش «کیدامبی» (۲۰۱۰) همسوست. نیز یاهو، لایکاس، پیکسرچ و اکسالید بر اساس معیارهای این پژوهش و درصد فراوانی کسب شده، به ترتیب رتبه‌های دوم تا پنجم را کسب کرده‌اند. این نکته نیز قابل توجه است که بر اساس نتایج حاصل از پژوهش حاضر، موتورهای جستجوی عمومی در این زمینه نسبت به موتورهای جستجوی تخصصی مورد مطالعه عملکرد بهتری داشتند. در خاتمه، پیشنهاد می‌شود برای توجه بیشتر به نمایه‌سازی و بازیابی تصاویر در موتورهای جستجو، طراحان امکاناتی از جمله: ترکیب همزمان چند رنگ، افزودن کد مخصوصی برای حق مؤلف تصاویر، اعمال همه مؤلفه‌های این پژوهش به صورت منو و زیر منوی در موتورهای جستجوی تصاویر در نظر گرفته شود. همچنین، برای پژوهشهای آتی نیز پیشنهاد می‌شود کاربرد یادگیری کاربران برای رفع شکاف معنایی تصاویر، کاربرد تصاویر در مدیریت دانش و دانش معنایی، مؤلفه‌های رابط کاربر در نظامهای بازیابی تصاویر بررسی شود.

منابع

- اسفندیاری‌مقدم، علیرضا و زهره بهاری‌موفق (۱۳۸۸). امکانات جستجو در فراموتورهای جستجو دروب: رویکردی مبتنی برسیاه‌وارسی، فصلنامه علوم و فناوری اطلاعات، دوره ۲۵، شماره ۲، ص ۲۶۵-۲۸۷.

1. Ozendi.

- پورسیستانی، پردیس (۱۳۸۹). ارزیابی فاکتورهای مؤثر در نمایه‌سازی و بازیابی تصویر براساس محتوا در حوزه فشرده JPEG، دانشگاه پیام‌نور، پایان‌نامه کارشناسی ارشد مهندسی کامپیوتر، دانشکده کامپیوتر و فناوری اطلاعات.
- عباس‌پور، جواد (۱۳۸۴). نمایه‌سازی تصاویر: چالش و رویکردها، مجله کتابداری، سال نهم، دفترچهل و چهارم، ص ۱۶۷-۱۷۷.
- فتحیان، محسن (۱۳۹۰). بازیابی محتوایی تصاویر بر مبنای یادگیری ماشین از طریق تعامل با کاربر، دانشگاه کردستان، دانشکده فنی و مهندسی، پایان‌نامه کارشناسی ارشد مهندسی کامپیوتر.
- فرخ‌زاد، بی‌بی‌فاطمه (۱۳۹۰). بررسی میزان دقت در بازیابی تصاویر همراه با متن با شیوه خوشه‌بندی نتایج جستجو از موتور جستجوی Yippy، دانشگاه الزهراء، پایان‌نامه کارشناسی ارشد علوم انسانی، دانشکده روانشناسی و علوم تربیتی.
- کرمی، افسانه (۱۳۸۵). نمایه‌سازی تصویر در پایگاه‌های اطلاعاتی، نما مجله الکترونیکی پژوهشگاه اطلاعات و مدارک علمی ایران، شماره سوم، دوره هشتم.
- کوشا، کیوان (۱۳۸۱). ابزارهای کاوش در اینترنت: اصول، مهارت‌ها و امکانات جستجو در وب. تهران، کتابدار.
- لاکدشتی، ابوالفضل (۱۳۸۸). نمایه‌سازی و بازیابی داده‌های تصویری مبتنی بر محتوا و معناشناسی بصری در پایگاه داده تصویر، دانشگاه آزاد اسلامی، واحد علوم و تحقیقات تهران، رساله دکتری رشته کامپیوتر.
- منصوری، زهرا (۱۳۸۷). بازیابی تصاویر رنگی با استفاده از دانش بافت و رنگ در ساختار درختی دودویی، پایان‌نامه کارشناسی ارشد گرایش هوش مصنوعی، دانشگاه صنعتی شریف: دانشکده مهندسی کامپیوتر.
- مهرآبادی، میترا (۱۳۸۵). مقایسه موتورهای کاوش عمومی و باز لحاظ نحوه نمایه‌سازی اطلاعات (با تأکید بر حوزه کتابداری و اطلاع‌رسانی)، دانشگاه تهران، دانشکده علوم تربیتی و روانشناسی، پایان‌نامه کارشناسی ارشد.
- Baxter, G. & Douglas, A. (1995). Image Indexing and Retrieval: Some Problems and Proposed Solutions, New Library World, V96.
- Chu, H. (2001). Research in Image Indexing and Retrieval as Reflected in the Literature, American society for Information Science And Technology, Vol.52.(12), pp. 1011-11018.
- Choi, Y. and Rasmussen, E.M. (2003). Searching for images: the analysis of users' queries for image retrieval in American history, Journal of the American Society for Information Science and Technology, Vol.54, No.6, pp.498-511.
- Diamant, E. (2007). Modeling human-like intelligent image processing: An information processing perspective and approach, Image Communication, Vol. 22, No. 6, pp. 583 - 590.

- Guopin, Q. & Jeremy, M.& Xunli, F. (2007). Visual guided navigation for image retrieval, *Pattern Recognition*, Vol. 40, No 6, pp. 1711 - 1721.
- Hassan, Ibrahim & Zhang, Jin (2001). Image search engine future analysis, *Online Information Review*, 25(2). pp. 103-114. [on-line].available at: <http://www.emerald-library-com/ft>. (accessed 20 April 2013)
- Hove, L-J. (2004). Improving image retrieval with a thesaurus for shapes, the VORTEX prototype. Master's thesis. Faculty of Social Sciences, Department Of information Science and Media Studies, Bergen University, Norway.
- Ink pen, Diana... [et al]. (2008). Clustering for photo retrieval at image CLEF 2008, [on-line]. Available: http://www.site.uottawa.ca/~diana/publicactions/UO_Photo_CLEF2008_v3.pdf. (accessed 25 June 2013)
- Kidambi, P. (2010). Human-Computer Integrated Approach towards Content Based Image Retrieval, Wright State University.
- Ozendi, M. (2010). Viewpoint Independent Image Classification and Retrieval, the Ohio State University. *Review of Image Search Engines* (2013). <http://www.tasi.ac.uk/resources/searchengines> (<http://www.jiscdigitalmedia.ac.uk/guide/review-of-image-search-engines>). (accessed 28 April 2013).
- Roberts, H. E. (2001). A picture is worth a thousand words: art indexing in electronic databases, *Journal of the American Society for Information Science and Technology* 52 (11), pp. 911-916.
- Smith, J. R. (2001). Quantitative assessment of image retrieval effectiveness, *Journal of the American Society for Information Science and Technology* 52 (11), pp. 969-979.
- Vassilieva, N.S. (2009). Content-based image retrieval methods, *Programming and Computing Software*, Vol. 35, No. 3, pp. 158 - 180.