

Assessing Rate of Matching Content Retrieval Conformity with Images and its Challenges in Scientific Database Articles

Adel Soleimani Nezhad*

PhD in Knowledge and Information Sciences; Assistant Professor;
Department of Knowledge and Information Science;
Shahid Bahonar University of Kerman; Kerman, Iran;
Email: Adels2004@yahoo.com

Fariborz Droudi

PhD in Knowledge and Information Sciences; Assistant Professor;
Iranian Research Institute for Information Science and Technology
(IranDoc); Tehran, Iran Email: doroudi@irandoc.ac.ir

Zynab Morteza Pour

M.A. in Knowledge and Information Science; Shahid Bahonar
University of Kerman; Kerman, Iran;
Email: Zeynabmortezaipoor@yahoo.com

Received: 10, Sep. 2018 | Accepted: 13, Feb. 2019

Abstract: The purpose of this research is to match the content retrieval with images in the articles of scientific databases. The research method is a correlation-based procedure and the statistical population of the present study is the scientific databases such as Science Direct, Scopus, Medline and Web of Science. Six papers were extracted from each base along with a separate issue. First, the title, abstract, problem statement and conclusion of the articles were inserted into the software Extreme Picture Finder, separately. Then, the images related to title, abstract, problem statement and conclusions of the articles were extracted from the relevant software. The images extracted were inserted into the Visual Similarity Duplicate Image Finder software. The results show that the highest level of content matching was found with the images in the WebAsiensis database and the least amount of match in the database of the SinjServer. There were also the most similarities between the titles, the problem statement, the abstract and the conclusions with the images of the articles in the bases.

Failure to observe the standards of visualization in the images used in scientific articles was a serious challenge. Images that have been removed from image editors (text, color, edges, margins, etc.) have not been recovered, and if recovered, they have had a fairly similar resemblance. The lack of comprehensiveness of the superficial

Iranian Journal of
**Information
Processing and
Management**

Iranian Research Institute
for Information Science and Technology
(IranDoc)

ISSN 2251-8223

eISSN 2251-8231

Indexed by SCOPUS, ISC, & LISTA

Vol. 35 | No. 1 | pp. 173-200

Autumn 2019

<https://doi.org/10.35050/IJPM010.2019.006>



* Corresponding Author

capabilities used was the next challenge. Another problem is the non-observance of or compliance with a standard style sheet in the layout of images in the articles. Some images have been brought to the articles regardless of the text. This item was found to be smaller in the Web site, but was seen in many other sites.

Keywords: Image Retrieval, Content Based Image Retrieval, Image Matching Content, Challenges of Matching Content with the Image ,Scientific Databases

بررسی میزان مطابقت بازیابی محتوا با تصویرها و چالش‌های آن در مقاله‌های پایگاه‌های اطلاعاتی علمی

عادل سلیمانی نژاد

دکتری؛ علم اطلاعات و دانش‌شناسی؛ استادیار؛ بخش
علم اطلاعات و دانش‌شناسی؛ دانشگاه شهید باهنر کرمان،
پدیده‌آور رابط Adels2004@yahoo.com

فریبرز درودی

دکتری؛ علم اطلاعات و دانش‌شناسی؛ استادیار؛
پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)؛
doroudi@irandoc.ac.ir

زینب مرتضی پور

کارشناسی ارشد؛ علم اطلاعات و دانش‌شناسی؛
دانشگاه شهید باهنر کرمان؛
Zeynabmortezaapoor@yahoo.com



دریافت: ۱۳۹۷/۰۶/۱۹ | پذیرش: ۱۳۹۷/۱۱/۲۴ | مقاله برای اصلاح به مدت یک ماه و ۱۸ روز نزد پدیدآوران بوده است.

چکیده: هدف از این تحقیق بررسی میزان مطابقت بازیابی محتوا با تصویرها و چالش‌های آن در مقاله‌های پایگاه‌های اطلاعاتی علمی است. روش پیمایشی توصیفی بوده و جامعه آماری آن پایگاه‌های اطلاعاتی علمی «ساینس دایرکت»، «اسکوپوس»، «مدلاین» و «وب آو ساینس» است. ابتدا عنوان، چکیده، بیان مسئله و نتیجه‌گیری مقاله‌ها به صورت جداگانه وارد نرم‌افزار Extreme Picture Finder شد. سپس، تصویرهای مرتبط با عنوان، چکیده، بیان مسئله و نتیجه‌گیری مقاله‌ها از نرم‌افزار مورد نظر استخراج گردید. تصویرهای استخراج‌شده وارد نرم‌افزار Visual Similarity Duplicate Image Finder گردید تا مطابقت تصویرهای استخراج‌شده از نرم‌افزار و تصویرهای مقاله‌ها انجام شود. نتایج نشان داد که بیشترین میزان مطابقت محتوا با تصویر در پایگاه «وب آو ساینس» و کمترین میزان مطابقت در پایگاه «ساینس دایرکت» وجود داشت. همچنین به ترتیب، بیشترین شباهت بین عنوان‌ها، بیان مسئله، چکیده و نتیجه‌گیری با تصویرهای مقاله‌ها در پایگاه‌ها وجود داشت.

عدم رعایت استانداردهای مصورسازی در تصویرهای به کاررفته در مقاله‌های پایگاه‌های علمی چالشی جدی است. تصاویری که از استانداردهای

نشریه علمی | رتبه بین‌المللی
پژوهشگاه علوم و فناوری اطلاعات ایران
(ایرانداک)

شاپا (چاپی) ۲۳۰۱-۸۲۲۳

شاپا (الکترونیکی) ۸۲۳۱-۲۳۰۱

نمایه در SCOPUS، ISC، LISTA و

jipm.irandoc.ac.ir

دوره ۳۵ | شماره ۱ | صص ۱۷۳-۲۰۰

پاییز ۱۳۹۸

<https://doi.org/10.35050/JIPM010.2019.006>



مصورسازی (متن، رنگ، لبه، حاشیه و ...) به دور بودند، بازیابی نشدند و در صورت بازیابی از میزان شباهت نازلی برخوردار بودند. عدم جامعیت قابلیت‌های نرم‌افزارهای مورد استفاده، چالش بعدی بود. مسئله دیگر عدم رعایت یا پیروی از یک شیوه‌نامه استاندارد در چیدمان تصویرها در مقاله‌هاست. برخی تصویرها بدون توجه به متن در مقاله‌ها درج شده‌اند. این مورد در پایگاه «وب‌آوساینس» کمتر، اما در سایر پایگاه‌های مورد بررسی بسیار مشاهده گردید. با توجه به ابداع روش‌های جدید بازیابی تصویر و محتوا، نتیجه این بررسی نشان می‌دهد که عدم شباهت بین تصویرهای بازیابی شده با محتوا در مقالات پایگاه‌های معتبر علمی چشمگیر و گمراه‌کننده است.

کلیدواژه‌ها: بازیابی تصویر، بازیابی محتوا، مقاله‌های علمی، چالش‌های بازیابی تصویر، پایگاه اطلاعات علمی

۱. مقدمه

تصویرها ابزاری قدرتمند برای انتقال اطلاعات محسوب می‌شوند و بیشتر در ترکیب با متن‌ها می‌توانند اطلاعاتی را که با کلمه‌ها قابل توصیف نیستند، منتقل کنند. یک تصویر می‌تواند هزاران بار با ارزش‌تر از کلمه‌ها باشد، اما گاه هزار توصیفگر نیز برای توصیف محتوای یک پایگاه اطلاعاتی تصویری کافی نیست. «خرفی، زیو و برناردی» در مقاله خود کاربردهای تصویر را بررسی کرده‌اند. به نظر آن‌ها متن نمی‌تواند در شرح یک رویداد، تصویرهای پزشکی، توصیف یک مکان و غیره به خوبی اطلاعات را منتقل نماید. استفاده از تصویر مناسب برای نمایش هرچه بهتر محصول بسیار مهم است (Kherfi, Ziou and Bernardi 2004). ماهیت چندبعدی بودن تصویرها باعث می‌شود که هر کاربر با توجه به نیاز خود، نگاه ویژه‌ای به یک تصویر داشته باشد و حتی با هدفی جدا از هدف خالق تصویر، آن را به کار گیرد. وجود یک تصویر همراه با متن می‌تواند به درک و انتقال مفاهیم موجود در آن کمک کند؛ در حالی که همان تصویر در متنی دیگر مفهوم دیگری را می‌تواند تداعی نماید (Hassan and Zhang 2001).

با رشد روزافزون اینترنت و ابزارهای تصویربرداری رقومی، اندازه پایگاه‌های داده تصویری به سرعت در حال بزرگ‌تر شدن است. در چنین شرایطی، نیاز شدیدی به ابزارها و روش‌های کارا برای جست‌وجوی تصویرهای دلخواه در پایگاه‌های داده بزرگ به وجود آمده است. در سامانه‌های بازیابی تصویر بر اساس محتوا سعی بر این است که تصویرها از پایگاه داده، که به تصویر پرس‌وجوی کاربر بیشترین شباهت را دارند، به‌عنوان

نتیجه جست‌وجو بازیابی شوند. بازیابی تصویر بر اساس محتوا به‌عنوان یک جایگزین کامل‌تر و دقیق‌تر برای نظام‌های بازیابی تصویر بر اساس متن به‌شمار می‌رود (Liu et al. 2007). برای اندازه‌گیری میزان شباهت تصاویر در اکثر سامانه‌های بازیابی بر اساس محتوا، ویژگی‌های سطح پایین تصویر از جمله رنگ، بافت و شکل به‌صورت خودکار از تصویرها استخراج می‌شود و بُرداری که بیانگر ویژگی‌های تصویر است، ساخته می‌شود. استخراج ویژگی، اساسی‌ترین قدم در ایجاد یک سامانه بازیابی تصویر بر اساس محتواست و نقش بسیار تعیین‌کننده‌ای در دقت سامانه بازیابی دارد. روش‌های مختلفی برای استخراج بُردارهای ویژگی از تصاویر موجود است؛ از جمله فیلترهای گابور و هیستوگرام راستای لبه برای استخراج اطلاعات فرکانسی و هیستوگرام رنگ برای استخراج اطلاعات مربوط به رنگ در تصاویر رنگی (Aujol et al. 2006).

با توجه به این‌که بیشتر اطلاعات علمی به‌صورت متنی چاپ می‌شوند و فهم و درک آن‌ها گاه بسیار سخت است، جایگزینی یا همراهی این محتوا با تصویر می‌تواند فهم و درک بهتری از مقاله‌های علمی را در مخاطبان ایجاد کند. با این کار می‌توان از ماهیت نحوی مقاله‌ها به ماهیت معنایی آن‌ها پی برد. مسئله اصلی این پژوهش آن است که در بازیابی مقاله‌های علمی از پایگاه‌های اطلاعاتی علمی اکثر مقاله‌های علمی فاقد تصویر یا به‌ندرت همراه با تصویر هستند و در بیشتر موارد شباهت یا رابطه منطقی بین محتوا و تصویرهای بازیابی‌شده بسیار ضعیف است. این مسئله باعث گمراهی و عدم اعتماد کاربران نسبت به نتایج بازیابی محتوا به همراه تصویر می‌شود. اکثر پژوهش‌های انجام‌شده به روش‌ها، فنون، و مدل‌های انطباق بیشتر دو مقوله محتوا و تصویر در مقالات، به‌خصوص در مقالات علمی پرداخته یا از آن‌ها رونمایی شده است. مسئله پژوهش این است که با توجه به رشد بالای مقالات در پایگاه‌های اطلاعاتی، بازیابی توأم محتوا و تصویر برای کاربران، آگاهی از میزان انطباق و چالش‌های موجود ضرورت دارد تا بررسی در این خصوص انجام شود.

هدف اصلی این پژوهش آگاهی از میزان مطابقت تصویرهای بازیابی‌شده با محتوای مقاله‌ها در پایگاه‌های اطلاعاتی علمی «اسکوپوس»، «ساینس دایرکت»، «مدلاین» و «وب‌آو‌ساینس» است. همچنین، هدف‌های فرعی این پژوهش آگاهی از میزان مطابقت اجزای مقاله‌ها (عنوان، چکیده، بیان مسئله و نتیجه‌گیری) با تصویرهای بازیابی‌شده و شناسایی چالش‌های بازیابی و تطابق محتوا با تصویرها در پایگاه‌های اطلاعاتی علمی مورد

بررسی است. از این رو، از نرم‌افزار Extreme Picture Finder به منظور بازیابی محتوا با تصویرها و از نرم‌افزار Visual Similarity Duplicate Image Finder جهت مطابقت محتوا با تصویرهای بازیابی شده استفاده می‌کنیم.

۲. پیشینه پژوهش

پس از جست‌وجو در پایگاه‌های منابع اطلاعاتی و بررسی پژوهش‌های انجام‌شده مشخص شد که مطالعات در زمینه بازیابی تصویر مبتنی بر محتوا، هم در ایران هم در خارج از کشور انجام نشده است. لذا، تا جایی که مقدور بود، پژوهش‌های مرتبط با موضوع پژوهش آورده شده است.

در مقاله (Unar, Wang and Zhang (2018 با عنوان «ترکیب اطلاعات متنی و تصویری با استفاده از روش کرنل جهت بازیابی تصویرها مبتنی بر محتوا»، هر منطقه از یک تصویر دارای ارزش و مهم است، زیرا می‌تواند دربرگیرنده یک مفهوم، متن یا تصویر باشد. تاکنون ویژگی‌های بصری یا برجسته‌های اجتماعی برای بازیابی تصویرهای مشابه مورد استفاده قرار می‌گرفتند. تولید برجسته‌ها و کلمات کلیدی متن و تصویر جاسازی شده در تصویرها همراه با ویژگی‌های بصری در سطح پایین است. این روش دو بردار ویژگی را برای هر تصویر و متن اختصاص می‌دهد و آن‌ها را با استفاده از روش «کرنل»^۱ همپوشانی می‌کند. این روش از سه روش جست‌وجوی پرس‌وجوی تصویر، کلمات کلیدی، و ترکیبی از هر دو پشتیبانی می‌کند. نتایج تجربی بر روی مجموعه داده‌های معیار نشان می‌دهد که ویژگی‌های متنی می‌تواند به عنوان ویژگی‌های بصری برای برنامه‌های بازیابی تصویرها مبتنی بر محتوا مؤثر باشد. این در حالی است که در پژوهش‌های قبلی صرفاً از کلمات کلیدی به عنوان ویژگی‌های بصری در بازیابی تصویرها مبتنی بر محتوا استفاده می‌شد. به کارگیری روش «کرنل» این قابلیت را به وجود آورد که علاوه بر کلمات کلیدی، ویژگی‌های متنی هم به عنوان ویژگی‌های بصری در بازیابی تصویرها مبتنی بر محتوا بسیار مؤثر واقع شوند.

در مقاله «آزودینیا و هاجو»، تحت عنوان «روش مبتنی بر بازخورد ارتباطی ترکیبی جدید برای بازیابی تصویر مبتنی بر محتوا»، به اهمیت استفاده از رویکردهای مؤثر برای

۱. «کرنل» Kernel یکی از روش‌های برآورد تابع احتمال است که تعمیمی از برآوردگر ساده مستطیلی که دو پارامتر تابع کرنل و پهنای باند را تعیین می‌کند.

بازیابی تصویرهای مرتبط، تصویر پرس‌وجو اشاره شده است که از بازیابی تصویر مبتنی بر محتوا برای حل این مسئله استفاده می‌شود. در این مقاله یک روش بازخورد ارتباطی جدید ارائه شده است که می‌تواند دقت سیستم‌های بازیابی مبتنی بر محتوا را بهبود بخشد. روش پیشنهادی بر مبنای بازخورد ارتباطی چند پرس‌وجو و مطابقت تابع تشابه است (Azodinia and Hajdu 2016). هر چند این روش تحول خوبی در جهت مطابقت بین محتوا و تصویر در بازیابی است، اما برآورد میزان مشابهت محتوا و تصویر در سطح بالایی نیست.

در پژوهشی با عنوان «روش هوش مصنوعی ترکیبی برای اطلاع از بازیابی تصویر مبتنی بر محتوا» که توسط «مشرم، تاکاری و گداده» انجام شد، یک رویکرد برای بازیابی تصویر مبتنی بر محتوا پیشنهاد شده که در آن تصویرهای پایگاه داده‌ها به دسته‌های بهینه‌سازی شده برای فرایند بازیابی بیشتر دسته‌بندی شده‌اند. گوناگونی الگوریتم‌های دسته‌بندی شده اجرا و نتایج مقایسه شدند. در میان تمام یافته‌ها الگوریتم ترکیبی (ای‌سی‌پی‌اس^۱) بهتر از الگوریتم‌های پایه نظیر الگوریتم (پی‌اس‌ا^۲) و (ای‌سی‌ا^۳) و (کی-مینز)^۴ و غیره هستند. در این مقاله با ترکیب الگوریتم‌های موجود ضمن دسته‌بندی تصویرهای نزدیک به هم و محتوای نزدیک به هم، دسته‌بندی‌های بیشتری ایجاد شد. نتیجه این ترکیب خروجی‌های بهتر نسبت به الگوریتم‌های قبل از ترکیب بود، که بهینه‌سازی بازیابی تصویرها بر اساس محتوا حاصل شد (Meshram, Thakare, and Gudadhe 2016).

«خوداسکر و لدهاک» در پژوهشی با عنوان «بازیابی پیشرفته تصویر با راهبرد طبقه‌بندی موضوعی»، از روش طبقه‌بندی موضوعی استفاده کردند. این تحقیق یک زمینه پیشرفته را بر اساس سیستم بازیابی تصویر ارائه می‌دهد که هر موضوع بر اساس راهبرد طبقه‌بندی، به یک تصویر اختصاص داده می‌شود تا عملکرد بازیابی به حد چشمگیری بهبود یابد. این مقاله همچنین، مرور کلی از طبقه‌بندی تصویر را ارائه داده است که گروه‌های مختلفی از طبقه‌بندی، عوامل مؤثر بر دقت طبقه‌بندی و کاربردهای حاضر طبقه‌بندی با تکنیک‌های پیشرفته را دربرمی‌گیرد. این راهبرد طبقه‌بندی ارائه شده از سه قانون کارآموزی استفاده

-
1. ACPSO= Ant Colony Particle Swarm Optimization (ترکیبی از دو الگوریتم ای‌سی‌ا با پی‌اس‌ا)
 2. ACO= Ant Colony Optimization (معروف به الگوریتم کلونی مورچه‌ها، پیدا کردن کوتاه‌ترین مسیر از بین مسیرها)
 3. PSO=Particle Swarm Optimization (یک الگوریتم جدید بر پایه قوانین بهینه‌سازی انبوه ذرات)
 4. K_means (الگوریتم‌های خوشه‌بندی داده‌ها در داده کاوی و متن کاوی)

می‌کند: سطح پایین، سطح بالا و قوانین تخصصی که دقت طبقه‌بندی و اثربخشی و در نهایت، دست‌درازی در کیفیت طبقه‌بندی را اصلاح می‌کند. نتیجه تجربی تکامل اجرایی در وضوح، دقت و زمان بازیابی را نشان می‌دهد (Khodaskar and Ladhake 2015).

«جنی، ماندلا و سانار» در مقاله «بازیابی تصویر بر اساس محتوا با استفاده از مقایسه رشته‌های رنگ»، رویکردی جدید برای بازیابی تصویر مبتنی بر محتوا بر اساس طبقه‌بندی پایگاه داده را با استفاده از دستگاه بردار پشتیبان و کدگذاری رشته رنگ برای انتخاب ویژگی ارائه کردند. به‌طور کلی، مشارکت در حوزه روبه‌رشد پایگاه داده تصویری فرایند کارآمدی برای به‌دست آوردن تصویرها از پایگاه داده است. با استفاده از طبقه‌بندی پایگاه داده‌ها می‌توان عملکرد بازیابی تصویر مبتنی بر محتوا نسبت به «سی‌بی‌آر»^۱ نرمال را بهبود بخشید؛ یعنی بدون طبقه‌بندی پایگاه داده. در نهایت، این طبقه‌بندی پایگاه داده و انتخاب ویژگی کدگذاری رشته رنگی نتایج بهتری ارائه می‌دهد (Jenni, Mandala, and Sunar 2015).

«کاا» و همکاران در پژوهشی دیگر با عنوان «بازیابی تصویر مبتنی بر محتوا با استفاده از هندسه اطلاعات با ابعاد بزرگ» یک روش جدید بازیابی تصویر مبتنی بر محتوا بر اساس نظریه اطلاعات با ابعاد بزرگ را ارائه می‌کنند. مزایای این رویکرد پیشنهادی آن است که شکاف معناشناختی در الگوریتم‌های بازیابی تصویر مبتنی بر محتوا را برطرف می‌کند. در این مقاله ابتدا، یک الگوریتم کیهانی زاویه بردار فضایی جدید چندبعدی از هندسه با ابعاد بزرگ ارائه شد. سپس، شرح مفصلی از روش بازیابی تصویرها، شامل پیشنهاد یک روش بلوک تصویر همپوشانی و تعریف یک درجه شباهت بین تصویرها در غیر زیرمجموعه‌های اطلاعاتی ارائه گردید (Cao et al. 2014).

«پوررضا» در پایان‌نامه خود یکی از مهم‌ترین چالش‌های بازیابی تصویرها را گسترش سریع تصویرها توسط انواع منابع تولیدکننده اطلاعات در اینترنت و شبکه‌های اجتماعی می‌داند که موجب کاهش کارایی این سیستم‌ها شده است. از این رو، در این پایان‌نامه، یک سیستم بازیابی تصویر تکراری-جزئی ارائه شده است که شامل سه گام اصلی است: تشخیص برجستگی، استخراج ویژگی و تعیین شباهت. در گام اول، بر اساس اصل رنگ پس‌زمینه، نقشه برجستگی تصویر تعیین می‌گردد. در گام دوم، نقاط کلیدی ناحیه برجسته

1. CBR= Content based retrieval (بازیابی مبتنی بر محتوا)

و ویژگی‌های آن‌ها با استفاده از الگوریتم SIFT^۱ محاسبه می‌شود، و برای ناحیه اطراف هر نقطه کلیدی، یک هیستوگرام رنگ به دست می‌آید. در گام سوم، با استفاده از ویژگی‌های SIFT محاسبه شده، تطابق هر جفت نقطه کلیدی از تصویر منتخب کاربر و تصویر پایگاه داده بررسی می‌شود. روش پیشنهادی بر روی پایگاه داده‌های IPDID و INSTRE پیاده‌سازی و ارزیابی شده و بر اساس معیار شباهت MAP^۲، به ترتیب، نتایج ۸۰ درصد و ۶۵ درصد حاصل گردید. نتایج پیاده‌سازی، حاکی از عملکرد مناسب روش پیشنهادی است (۱۳۹۴).

«شمسی گوشکی» و همکاران در مقاله «روشی برای بازخورد ربط بر اساس بهبود تابع شباهت در بازیابی تصویر بر اساس محتوا»، به بررسی روشی برای بازخورد بر اساس تابع شباهت در بازیابی تصویر بر اساس محتوا می‌پردازند. در این روش وزن هر نوع ویژگی با توجه به رتبه تصویرهای مرتبط در بازیابی بر اساس آن نوع ویژگی تنظیم و برای تصحیح وزن هر مؤلفه ویژگی از میانگین و انحراف معیار آن مؤلفه روی تصویرهای مرتبط و نامرتب استفاده می‌شود. روش پیشنهادی روی یک پایگاه تصویر شامل ده هزار تصویر از ۸۲ گروه معنایی متفاوت آزموده شده است. نتایج آزمایش برتری روش پیشنهادی را نسبت به روش‌های موجود نشان می‌دهد (شمسی و دیگران ۱۳۹۳).

همچنین، «بیات» و همکاران او در مقاله «بازیابی مبتنی بر محتوای تصویرها با استخراج ویژگی‌های بافت و رنگ»، تصویر با استفاده از مورفولوژی و هیستوگرام رنگی روشی برای استخراج ویژگی‌های بافت و رنگ تصویرها به منظور بازیابی مبتنی بر محتوای تصویر ارائه کردند که در آن با توجه به خصوصیات دو روش عناصر ساختاری و هیستوگرام رنگی، استخراج ویژگی تصویرها در دو مرحله انجام می‌شود. در مرحله اول، استخراج و دستیابی به ویژگی بافت تصویر و در مرحله دوم، از هیستوگرام رنگی برای استخراج ویژگی رنگ تصویر استفاده می‌شود. نتایج تجربی نشان می‌دهد که سیستم بازیابی تصویر ارائه شده در این مقاله، به دلیل استفاده همزمان از ویژگی بافت و رنگ تصویر از کارایی بالایی برخوردار است (۱۳۹۳).

«ذوالفقاری و خسروی» در مقاله خود با عنوان «روشی سریع در بازیابی تصویرها مبتنی

1. SIFT= Scale Invariant Feature Transform (الگوریتم تشخیص الگو در تصاویر نوری)

2. MAP=Mean Average Precision (معیاری است برای محاسبه میانگین متوسط دقت شباهت)

بر محتوا با استفاده از ترکیب ویژگی لبه و رنگ^۱ رویکرد جدیدی برای بازیابی تصویرها مبتنی بر محتوا ارائه کرده‌اند (۱۳۹۲). جهت استخراج ویژگی از دو ویژگی رنگ^۱ HSV و ویژگی گرادینان^۲ لبه استفاده شده است. بر اساس آزمایش‌های انجام شده بر روی پایگاه داده Wang^۳ مشتمل بر ۱۰۰۰ تصویر و ۱۰ کلاس معنایی، مشاهده می‌شود که دقت نرخ بازیابی در مقایسه کارهای پیشین ارتقا پیدا کرده است.

«حیاتی، سریزدی و نظام‌آبادی‌پور» در پژوهشی با عنوان «جداسازی تصویر به مؤلفه‌های بافت و ساختار، برای بازیابی تصویر بر اساس محتوا» راهکار جدیدی برای بهبود روش‌های استخراج ویژگی در سیستم بازیابی تصویر بر اساس محتوا با استفاده از جداسازی تصویر به مؤلفه‌های بافت و ساختار ارائه کردند (۱۳۹۱). جداسازی تصاویر به بافت و ساختار و استخراج ویژگی از هر بخش برای ویژگی‌های هیستوگرام لبه، هیستوگرام رنگ و گابور مورد بررسی قرار گرفت. در این مقاله روش جدیدی برای استخراج ویژگی‌های سطح پایین و نمایه‌سازی تصویر بر مبنای جداسازی تصویر به مؤلفه‌های بافت و ساختار ارائه شده است. این روش جداسازی جدید در مقایسه با روش‌های^۴ ROF از سرعت و دقت بیشتری برخوردار است. در نتیجه، استفاده همزمان از همه اطلاعات فرکانسی تصاویر بافت و ساختار، بردار ویژگی جدیدی با ترکیب بردارهای ویژگی گابور بخش بافت و ساختار جهت بازیابی پیشنهاد می‌شود که باعث افزایش دقت قابل ملاحظه‌ای در بازیابی تصویر می‌شود.

بررسی نتایج پژوهش‌های انجام شده نشان می‌دهد که در پژوهش «یونار، وانگ و ژانگ» رویکردی جدید برای بازیابی تصویرهای متنی مشابه با بهره‌برداری از ویژگی‌های بصری و متنی از تصویر حاصل شد. روش «کرنل»، از سه روش جست‌وجوی پرس‌وجوی تصویر، کلمات کلیدی و ترکیبی از متن و تصویر پشتیبانی می‌کند. این پژوهش با ارائه این رویکرد، روشی جدید در بازیابی تصویر توأم با محتوا را معرفی کرد که تحول مهمی بود، اما تطبیق محتوا و تصویر در شباهت عناصر از اهداف این پژوهش نبوده است (Unar,).

۱. HSV= Hue, (رنگدانه), Saturation (اشباع), Value (روشنایی)

۲. Gradient

۳. یک روش جدا سازی تصویر است و از سرنام‌های ابداع کنندگان آن VESE و OSHAR اختصار درست شده است.

۴. روشی جهت حذف نویز تصویر است که از سرنام‌های ابداع کنندگان آن RUDIN, OSHAR, FATEMI اختصار درست شده است.

(Wang, and Zhang 2018). اگر بازیابی را مرحله اول بدانیم، تطبیق نتایج مرحله بعدی است که ما در این پژوهش به دنبال آن هستیم. «آزودینیا و هاجو» در مقاله خود با استفاده از فن بازخورد ارتباطی، تصویرهای مربوطه را در پاسخ به یک پرس‌وجو بازیابی می‌کنند که می‌تواند دقت سیستم‌های بازیابی مبتنی بر محتوا را بهبود بخشد. فن بازخورد به کاررفته در این پژوهش میزان ارتباط تصاویر با یکدیگر را مورد سنجش قرار می‌دهد و اما قادر به سنجش ارتباط تصویر با محتوا نیست. هر چند این پژوهش در بازیابی تصویر فن جدیدی را معرفی می‌کند، اما به بازخورد شباهت بین تصویر و محتوا نمی‌پردازد (Azodinia and Hajdu 2016). «مشرم، تاکاری و گداده» ثابت کردند که الگوریتم ترکیبی (ای‌سی‌پی‌اس‌ا) بهتر از الگوریتم‌های پایه نظیر الگوریتم «پی‌اس‌ا» و «ای‌سی‌ا» و «کا-مینز» در بازیابی تصویرهای مبتنی بر محتوا می‌تواند مؤثر واقع شود، اما به هیچ معیار سنجش میزان شباهت بین تصویرهای بازیابی شده با محتوا اشاره ندارند (Meshram, Thakare, and Gudadhe 2016). پژوهش «جنی، ماندلا و سانار» نشان می‌دهد که طبقه‌بندی پایگاه داده و انتخاب ویژگی کدگذاری رشته رنگی نتایج بهتری ارائه می‌دهد. این پژوهش به ساختار محتوایی پایگاه‌های داده و ویژگی‌های کدگذاری می‌پردازد و جنبه محتوا را مورد بحث قرار نمی‌دهد (Jenni, Mandala and Sunar 2015). «کا» و همکاران با ارائه یک روش جدید بازیابی تصویر مبتنی بر محتوا بر اساس نظریه اطلاعات، الگوریتم کیهانی زاویه بردار فضایی جدید چند بعدی را از هندسه با ابعاد بزرگ ارائه کردند. بررسی این پژوهش نشان می‌دهد که آن‌ها به دنبال یافتن یک الگوریتم جدید در بازیابی تصویر مبتنی بر محتوا هستند، ولی تطبیق تصویر با محتوا را دنبال نمی‌کنند (Cao et al. 2014). پژوهش «شمسی گوشکی» و همکاران نشان داد که روش بازخورد ربط بر اساس بهبود تابع شباهت در بازیابی تصویر بر اساس محتوا نسبت به روش‌های موجود برتری دارد. در این پژوهش بررسی شباهت بین تصویرهای بازیابی شده با محتوا برای پژوهشی‌های آینده پیشنهاد می‌شود. پژوهش «بیات» و همکاران نشان می‌دهد که استفاده همزمان از ویژگی‌های بافت و رنگ در نظام بازیابی تصویر، باعث افزایش دقت در بازیابی تصاویر می‌شود (۱۳۹۳). این پژوهش راهکارهای مناسبی را برای بازیابی تصویر ارائه می‌کند و بیشترین تأکید آن بر ویژگی انحصاری تصویر است تا محتوا. پژوهش «ذوالفقاری و خسروی» نشان داد که روش مورد استفاده آن‌ها در بازیابی تصویرها، دقت بازیابی را در مقایسه با کارهای پیشین ارتقا داده است. این پژوهش نیز به بازیابی تصویر و دقت در

بازیابی اشاره دارد (۱۳۹۲). «حیاتی، سریزدی و نظام‌آبادی‌پور» در پژوهش خود با ارائه یک الگوریتم جدید، روش‌های استخراج ویژگی در سیستم بازیابی تصویر بر اساس محتوا را بهبود بخشیدند (۱۳۹۱). پیشینه‌های حاضر از مرتبط‌ترین پژوهش‌های موجود هستند که از بین پژوهش‌های زیادی انتخاب شده‌اند. همان‌طور که در تحلیل پیشینه‌ها اشاره شد، هر پیشینه‌ای به جنبه‌ای خاص از بازیابی تصویر یا محتوا پرداخته است. بیشترین مورد بررسی مربوط به راهکارهای جدید برای جست‌وجوی تصاویرها و محتوا به صورت جداگانه یا توأم است. آنچه این پژوهش را از پژوهش‌های مذکور متمایز می‌سازد، مطابقت و درجه تشابه بین تصویرهای بازیابی شده با محتوا و تصویرهای مقاله‌ها در پایگاه‌های علمی است. در این پژوهش، شناسایی نرم‌افزارهای بازیابی تصویرها و مطابقت تصویر با متن بر اساس ویژگی‌ها و رویکردهای جدید اهمیت دارد.

۳. روش پژوهش

پژوهش حاضر از نوع کاربردی و با روش پیمایشی توصیفی انجام شده است. در این پژوهش نرم‌افزارهای متعددی مورد بررسی قرار گرفته و در نهایت و بعد از آزمون و خطای بسیار نرم‌افزارهای مناسب این پژوهش انتخاب شدند. در این پژوهش با استفاده از نرم‌افزار Extreme Picture Finder ابتدا مقاله‌های مورد نظر به صورت تک‌تک وارد نرم‌افزار شدند و سپس، فیلترهای مورد نیاز از قبیل تعداد دانلودهای تصاویرها، آدرس، تاریخ (۲۰۱۲-۲۰۱۷) به جهت آخرین سال‌های انتشار مقاله‌ها و با توجه به کار زیاد تحلیل و مقایسه بازه زمانی پنج‌ساله انتخاب شد و سایر فیلترها مثل حجم، بافت و اندازه‌های تصاویرها اعمال گردید. سپس، عمل دانلود از نرم‌افزار مربوطه انجام و تصویرهای دانلودشده توسط نرم‌افزار در یک پوشه ذخیره شد تا انطباق با تصویرها و محتوا انجام شود. عمل جست‌وجو در این نرم‌افزار با استفاده از عنوان، چکیده، بیان مسئله و نتیجه‌گیری مقاله‌ها از پایگاه‌های اطلاعاتی علمی «ساینس دایرکت»^۱، «اسکوپوس»^۲، «وب‌آوساینس»^۳ و «مدلاین»^۴ با توجه به پوشش‌های موضوعی گسترده و اعتبار بالا انجام شد. با ایجاد یک پوشه فرعی برای هر

1. Science Direct (<https://www.sciencedirect.com>)

2. Scopus (<https://www.scopus.com>)

3. Web Of Science (<https://clarivate.com/products/web-of-science>)

4. Medline (<https://www.medline.com>)

موتور جست‌وجو و ایجاد یک پوشه فرعی برای هر عبارت جست‌وجو در یک پنجره جدید یک عنوان و یک فهرست برای عبارت مورد جست‌وجو جدید اختصاص داده شد.

جامعه آماری پژوهش حاضر پایگاه‌های اطلاعاتی علمی «ساینس دایرکت»، «اسکوپوس»، «وب‌آوساینس» و «مدلاین» هستند. در این پژوهش تعدادی مقاله از هر یک از پایگاه‌های اطلاعاتی بین سال‌های ۲۰۱۲ تا ۲۰۱۷ و بر اساس پوشش موضوعی انتخاب شد. در پایگاه اطلاعاتی «ساینس دایرکت» مقاله‌ها بر اساس موضوع شیمی، در پایگاه اطلاعاتی «وب‌آوساینس» مقاله‌ها بر اساس موضوع معماری، در «اسکوپوس» مقاله‌ها بر اساس موضوع رایانه، و در «مدلاین» مقاله‌ها بر اساس موضوع پزشکی انتخاب شدند و سپس، به‌صورت دستی شش مقاله، که دارای تصویر بودند، از بین سایر مقاله‌ها استخراج گردید. با استفاده از نرم‌افزارها Extreme Picture Finder و Visual Similarity Duplicate Image Finder عمل جست‌وجو و تجزیه و تحلیل انجام شد. عمل انطباق محتوای استخراج‌شده با تصویرهای موجود در پایگاه‌های «ساینس دایرکت»، «اسکوپوس»، «وب‌آوساینس» و «مدلاین» بر روی آن‌ها انجام و نتیجه انطباق استخراج گردید.

این پژوهش به دنبال پاسخگویی به پرسش‌های زیر است:

- ◇ میزان مطابقت محتوا با تصویرها در پایگاه‌های اطلاعاتی علمی چقدر است؟
- ◇ میزان مطابقت تصویرها با عنوان‌های مقاله‌ها در پایگاه‌های اطلاعاتی علمی چقدر است؟
- ◇ میزان مطابقت تصویرها با چکیده مقاله‌ها در پایگاه‌های اطلاعاتی علمی چقدر است؟
- ◇ میزان مطابقت تصویرها با بیان مسئله مقاله‌ها در پایگاه‌های اطلاعاتی علمی چقدر است؟
- ◇ میزان مطابقت تصویرها با نتیجه‌گیری مقاله‌ها در پایگاه‌های اطلاعاتی علمی چقدر است؟
- ◇ شناسایی چالش‌های بازاریابی و تطابق محتوا با تصویرها در پایگاه‌های اطلاعاتی علمی چقدر است؟

۴. یافته‌های پژوهش

در هر یک از جدول‌های توصیف پایگاه عنوان مقاله‌ها، تعداد تصویرهای موجود در

هر مقاله، تعداد تصویرهای باز یابی شده از عنوان، چکیده، بیان مسئله و نتیجه گیری مقاله‌ها توسط نرم‌افزار، تعداد تصویرهای مرتبط با مقاله‌ها، تعداد تصویرهای غیر مرتبط با مقاله‌ها و درصد شباهت بین تصویرهای مقاله‌ها و تصویرهای دانلود شده توسط نرم‌افزار آورده شده است. در این جدول‌ها، تعداد شش مقاله به صورت تصادفی در موضوعات پزشکی، شیمی، معماری و کامپیوتر از پایگاه‌های «مدلاین»، «ساینس دایرکت»، «وب‌آوساینس» و «اسکوپوس» استخراج شد و تعداد تصویرهای موجود در هر مقاله به صورت بصری شمارش گردید. در ادامه، با استفاده از نرم‌افزار Etreme Picture Finder عنوان‌ها، چکیده، بیان مسئله و نتیجه گیری مقاله‌ها استخراج و در محیط نرم‌افزار وارد شدند تا تصویرهای مرتبط با عنوان‌ها، چکیده، بیان مسئله و نتیجه گیری مقاله‌های استخراج شده باز یابی شوند. در نهایت، نتایج خروجی با عنوان تعداد تصویرهایی باز یابی شده با استفاده از نرم‌افزار مذکور استخراج گردید. برای حصول اطمینان از ارتباط تصویرهای باز یابی شده با مقاله‌ها توسط نرم‌افزار عمل تطبیق به صورت بصری انجام شد و نتایج در ستون تصویرهای مرتبط آورده شد. با توجه به نیاز به دقت و سرعت و درصد تشابه بیشتر بین تصویرهای موجود و تصویرهای باز یابی شده در اینجا از نرم‌افزار Visual Similarity Duplicate Image Finder با درصد شباهت ۵۰ درصد استفاده شد که نتایج در ستون درصد شباهت بین تصویرهای مقاله‌ها و تصویرهای باز یابی شده توسط نرم‌افزار آورده شده است. پس از وارد کردن عنوان‌ها به محیط نرم‌افزار Etreme Picture Finder برای استخراج تصویرها، عنوان مقاله‌ها به صورت کامل، ولی چکیده، بیان مسئله و نتیجه گیری مقاله‌ها به دلیل طولانی بودن و حجیم بودن، مطالبی که بیشترین مفهوم را می‌رساند انتخاب و وارد نرم‌افزار شد.

۴-۱. توصیف پایگاه «مدلاین»

جدول ۱. توصیف تعداد و درصد شباهت تصویرهای موجود در بخش‌های (عنوان، چکیده، بیان مسئله و نتیجه‌گیری) مقاله‌های پایگاه مدلاین با تصویرهای بازیابی شده از نرم‌افزار

| ردیف | عنوان مقاله | تعداد تصویرهای موجود در مقاله‌ها | عنوان | تعداد تصویرهای مرتبط | چکیده | تعداد تصویرهای بازیابی شده | بیان مسئله | تعداد تصویرهای بازیابی شده | نتیجه‌گیری | تعداد تصویرهای مرتبط |
|------|---|----------------------------------|---|----------------------|----------------------------|----------------------------|----------------------------|----------------------------|-----------------------|----------------------|
| ۱ | Dissecting the Role of the Extracellular Matrix in Heart Disease: Lessons from the Drosophila Genetic Model | ۷ | Efficacy of Oxycodone for Dyspnea in End-stage Heart Failure with Renal Insufficiency: A Case Report | ۴ | تعداد تصویرهای بازیابی شده | ۱۷ | تعداد تصویرهای بازیابی شده | ۲۵ | ۱۰ تصویر بین ۵۱ تا ۸۹ | ۲ |
| ۲ | Efficacy of Oxycodone for Dyspnea in End-stage Heart Failure with Renal Insufficiency: A Case Report | ۴ | Impact of Labile Zinc on Heart Function From Physiology to Pathophysiology | ۲۲ | تعداد تصویرهای بازیابی شده | ۳۵ | تعداد تصویرهای بازیابی شده | ۲۲ | ۷ | ۲۲ |
| ۳ | Impact of Labile Zinc on Heart Function From Physiology to Pathophysiology | ۲ | Profile of sacubitril/valsartan in the treatment of heart failure: patient selection and perspectives | ۱۶ | تعداد تصویرهای بازیابی شده | ۲۲ | تعداد تصویرهای بازیابی شده | ۲۲ | ۲ | ۲۲ |
| ۴ | Profile of sacubitril/valsartan in the treatment of heart failure: patient selection and perspectives | ۱ | تعداد تصویرهای بازیابی شده | ۲۲ | تعداد تصویرهای بازیابی شده | ۲۲ | تعداد تصویرهای بازیابی شده | ۲۲ | ۲ | ۲۲ |
| ۵ | تعداد تصویرهای بازیابی شده | ۲۲ | تعداد تصویرهای بازیابی شده | ۲۲ | تعداد تصویرهای بازیابی شده | ۲۲ | تعداد تصویرهای بازیابی شده | ۲۲ | ۲ | ۲۲ |
| ۶ | تعداد تصویرهای بازیابی شده | ۲۲ | تعداد تصویرهای بازیابی شده | ۲۲ | تعداد تصویرهای بازیابی شده | ۲۲ | تعداد تصویرهای بازیابی شده | ۲۲ | ۲ | ۲۲ |
| ۷ | تعداد تصویرهای بازیابی شده | ۲۲ | تعداد تصویرهای بازیابی شده | ۲۲ | تعداد تصویرهای بازیابی شده | ۲۲ | تعداد تصویرهای بازیابی شده | ۲۲ | ۲ | ۲۲ |
| ۸ | تعداد تصویرهای بازیابی شده | ۲۲ | تعداد تصویرهای بازیابی شده | ۲۲ | تعداد تصویرهای بازیابی شده | ۲۲ | تعداد تصویرهای بازیابی شده | ۲۲ | ۲ | ۲۲ |
| ۹ | تعداد تصویرهای بازیابی شده | ۲۲ | تعداد تصویرهای بازیابی شده | ۲۲ | تعداد تصویرهای بازیابی شده | ۲۲ | تعداد تصویرهای بازیابی شده | ۲۲ | ۲ | ۲۲ |
| ۱۰ | تعداد تصویرهای بازیابی شده | ۲۲ | تعداد تصویرهای بازیابی شده | ۲۲ | تعداد تصویرهای بازیابی شده | ۲۲ | تعداد تصویرهای بازیابی شده | ۲۲ | ۲ | ۲۲ |

| ردیف | عنوان مقاله‌ها | تعداد تصویرهای موجود در مقاله‌ها | عنوان | چکیده | بیان مسئله | نتیجه گیری |
|----------------|--|----------------------------------|---------------------------|------------------------------------|---------------------------|--|
| ۵ | Knitting for heart valve tissue engineering | ۱۸ | تعداد تصویرهای بازایی شده | تعداد تصویرهای بازایی شده مقاله‌ها | تعداد تصویرهای بازایی شده | تعداد شباهت بین تصویرهای مقاله‌ها و تعداد درصد شباهت بین تصویرها |
| ۶ | Centile Curves for Velocity-Time Integral Times Heart Rate as a Function of Ventricular Length The Use of Minute Distance Is Advantageous to Enhance Clinical Reliability in | ۷ | تعداد تصویرهای مرتبط | تعداد تصویرهای بازایی شده | تعداد تصویرهای بازایی شده | تعداد شباهت بین تصویرهای مقاله‌ها و تعداد درصد شباهت بین تصویرها |
| جمع | | ۴۱ | ۲۲۰ | ۱۴۹ | ۱۴ | ۹۹ |
| میانگین (درصد) | | ۴۱/۰۸ | ۲۵/۳۳ | ۲۵/۳۳ | ۲۵ | ۲۴ |

با توجه به نتایج جدول ۱، از شش مقاله مورد جست‌وجو در پایگاه «مدلاین»، ۴۱ تصویر به‌روش بصری از مقاله‌ها استخراج شد. با استفاده از نرم‌افزار Extreme Picture Finder، از ۲۲۰ تصویر بازایی شده توسط نرم‌افزار، ۱۴۹ مورد با عنوان‌های مقاله‌ها مرتبط بودند. همچنین، جهت بررسی میزان شباهت بین تصویرهای موجود در مقاله‌ها و تصویرهای بازایی شده توسط نرم‌افزار از درصد شباهت ۵۰ درصدی استفاده شده است و در کل، میانگین درصد شباهت بین تصویرهای موجود در مقاله‌ها و تصویرهای بازایی شده توسط نرم‌افزار ۴۱/۰۸ درصد است. از تعداد ۱۶۸ تصویر بازایی شده از مقاله‌ها توسط نرم‌افزار، ۹۹ مورد با چکیده مقاله‌ها مرتبط بودند. با بررسی میزان شباهت حداقل ۵۰ درصدی

بین تصویرهای موجود در مقاله‌ها و تصویرهای بازیابی شده توسط نرم‌افزار ۲۵/۳۳ درصد شباهت وجود دارد. از تعداد ۲۲۹ تصویر بازیابی شده از مقاله‌ها توسط نرم‌افزار، ۱۴۰ مورد با بیان مسئله مقاله‌ها مرتبط بودند. با بررسی میزان شباهت ۵۰ درصد بین تصویرهای موجود در مقاله‌ها و تصویرهای بازیابی شده توسط نرم‌افزار ۱۸/۳۳ درصد شباهت وجود دارد. از تعداد ۱۵۹ تصویر بازیابی شده از مقاله‌ها توسط نرم‌افزار، ۶۳ مورد با نتیجه‌گیری مقاله‌ها مرتبط بودند. با بررسی میزان شباهت ۵۰ درصد بین تصویرهای موجود در مقاله‌ها و تصویرهای بازیابی شده توسط نرم‌افزار ۲۱/۴۲ درصد شباهت وجود دارد. در نتیجه، در پایگاه اطلاعاتی «مدلاین» بیشترین درصد شباهت تصویرهای بازیابی شده توسط نرم‌افزار با عنوان مقاله‌ها و کمترین درصد شباهت با بیان مسئله بوده است.

۴-۲. توصیف پایگاه «ساینس دایرکت»

جدول ۲. توصیف تعداد و درصد شباهت تصویرهای موجود در بخش‌های (عنوان، چکیده، بیان مسئله و نتیجه‌گیری) مقاله‌های پایگاه «ساینس دایرکت» با تصویرهای بازیابی شده از نرم‌افزار

| ردیف | عنوان مقاله‌ها | تعداد تصویرهای موجود در مقاله‌ها | عنوان | چکیده | بیان مسئله | نتیجه‌گیری |
|------|--|----------------------------------|--------------------------------|----------------------------|----------------------------|--------------------------------|
| ۱ | Contrasted spatial and long-term trends in precipitation chemistry and deposition fluxes at rural stations in France | ۸ | تعداد تصویرهای مرتبط | تعداد تصویرهای بازیابی شده | تعداد تصویرهای بازیابی شده | تعداد و درصد شباهت بین تصویرها |
| | | ۲۴ | ۱۷ | ۲۵ | ۲۵ | ۰ |
| ۲ | Coordination chemistry of tris(azoly) phosphines | ۱۴ | تعداد و درصد شباهت بین تصویرها | تعداد تصویرهای مرتبط | تعداد تصویرهای مرتبط | تعداد تصویرهای مرتبط |
| | | ۲۹ | ۰ | ۲۹ | ۲۹ | ۲۹ |

| ردیف | عنوان مقاله‌ها | تعداد تصویرهای موجود در مقاله‌ها | | عنوان | چکیده | بیان مسئله | نتیجه‌گیری |
|----------------|---|----------------------------------|-------------------------|--------------------------------|----------------------|-------------------------|----------------------|
| | | تعداد تصویرهای مرتبط | تعداد تصویرهای بازمانده | | | | |
| ۴ | Native protein hydrogels by dynamic boronic acid chemistry | ۵ | ۰ | تعداد و درصد شباهت بین تصویرها | تعداد تصویرهای مرتبط | تعداد تصویرهای بازمانده | تعداد تصویرهای مرتبط |
| ۴ | The crystal chemistry and the compressibility of silicate-carbonate minerals: Spurrite, galuskinite and tilleyite | ۸ | ۰ | تعداد تصویرهای مرتبط | تعداد تصویرهای مرتبط | تعداد تصویرهای بازمانده | تعداد تصویرهای مرتبط |
| ۵ | Temporal and spatial patterns in the chemistry of wet deposition in Southern Alps | ۸ | ۰ | تعداد تصویرهای مرتبط | تعداد تصویرهای مرتبط | تعداد تصویرهای بازمانده | تعداد تصویرهای مرتبط |
| ۶ | Validation of Turbulence/Chemistry Interaction Models for use in Oxygen Enhanced Combustion | ۶ | ۰ | تعداد تصویرهای مرتبط | تعداد تصویرهای مرتبط | تعداد تصویرهای بازمانده | تعداد تصویرهای مرتبط |
| جمع | | ۴۹ | ۰ | ۱۲۲ | ۱۶۰ | ۱۰۰ | ۱۳۳ |
| میانگین (درصد) | | ۱۱/۱۶ | ۰ | ۱۷/۲۵ | ۱۵۲ | ۸۶ | ۱۰۳ |

با توجه به نتایج جدول ۲، از شش مقاله مورد جست‌وجو در پایگاه «ساینس دایرکت»، ۴۹ تصویر به روش بصری از مقاله‌ها استخراج شد. با استفاده از نرم‌افزار Extreme Picture Finder، از ۱۵۷ تصویر بازمانده توسط نرم‌افزار، ۱۲۲ مورد با عنوان‌های مقاله‌ها مرتبط بودند. همچنین، جهت بررسی میزان شباهت بین تصویرهای موجود در

مقاله‌ها و تصویرهای بازیابی‌شده توسط نرم‌افزار از شباهت ۵۰ درصدی استفاده شده است و در کل، میانگین درصد شباهت بین تصویرهای موجود در مقاله‌ها و تصویرهای بازیابی‌شده توسط نرم‌افزار ۱۱/۱۶ است. از تعداد ۱۶۰ تصویر بازیابی‌شده از مقاله‌ها توسط نرم‌افزار، ۱۰۰ مورد با چکیده مقاله‌ها مرتبط بودند. با بررسی میزان شباهت ۵۰ درصد، بین تصویرهای موجود در مقاله‌ها و تصویرهای بازیابی‌شده توسط نرم‌افزار هیچ شباهتی وجود ندارد. از تعداد ۱۵۲ تصویر بازیابی‌شده از مقاله‌ها توسط نرم‌افزار، ۸۶ مورد با بیان مسئله مقاله‌ها مرتبط بودند. با بررسی میزان شباهت ۵۰ درصد، بین تصویرهای موجود در مقاله‌ها و تصویرهای بازیابی‌شده توسط نرم‌افزار هیچ شباهتی وجود ندارد. از تعداد ۱۳۳ تصویر بازیابی‌شده از مقاله‌ها توسط نرم‌افزار، ۱۰۳ مورد با نتیجه‌گیری مقاله‌ها مرتبط بودند. با بررسی میزان شباهت ۵۰ درصد بین تصویرهای موجود در مقاله‌ها و تصویرهای بازیابی‌شده توسط نرم‌افزار ۱۷/۲۵ درصد شباهت وجود دارد. در نتیجه، در پایگاه اطلاعاتی «ساینس دایرکت» بیشترین درصد شباهت تصویرهای بازیابی‌شده توسط نرم‌افزار با قسمت نتیجه‌گیری مقاله‌ها، و کمترین درصد شباهت با چکیده و بیان مسئله بوده است.

| ردیف | عنوان مقاله‌ها | تعداد تصویرهای موجود در مقاله‌ها | عنوان | تعداد تصویرهای بازیابی شده | چکیده | بیان مسئله | نتیجه‌گیری |
|----------------|---|----------------------------------|----------------------------------|----------------------------|----------------------------|--------------------------------|----------------------|
| ۵ | Practical Arithmetic in Islamic Architecture: A Critical History and Survey | ۵ | تعداد تصویرهای موجود در مقاله‌ها | ۵ | تعداد تصویرهای بازیابی شده | تعداد و درصد شباهت بین تصویرها | تعداد تصویرهای مرتبط |
| ۶ | The Salman Mosque: Achmad Noe man's Critique of Indonesian Conventional Mosque Architecture | ۶ | تعداد تصویرهای مرتبط | ۶ | تعداد تصویرهای بازیابی شده | تعداد و درصد شباهت بین تصویرها | تعداد تصویرهای مرتبط |
| جمع | | ۷۵ | ۱۸ | ۱۲۳ | ۱۸ | ۸۴ | ۳۶ |
| میانگین (درصد) | | ۵۸/۲۵ | ۴۶/۱۶ | ۱۶۷ | ۵۷ | ۱۸۸ | ۱۰۷ |
| | | ۲ | ۱۱ | ۱۴ | ۲۵ | ۲۸ | ۲۹ |
| | | ۳ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۴ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۵ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۶ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۷ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۸ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۹ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۱۰ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۱۱ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۱۲ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۱۳ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۱۴ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۱۵ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۱۶ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۱۷ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۱۸ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۱۹ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۲۰ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۲۱ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۲۲ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۲۳ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۲۴ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۲۵ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۲۶ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۲۷ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۲۸ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۲۹ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۳۰ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۳۱ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۳۲ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۳۳ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۳۴ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۳۵ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۳۶ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۳۷ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۳۸ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۳۹ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۴۰ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۴۱ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۴۲ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۴۳ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۴۴ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۴۵ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۴۶ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۴۷ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۴۸ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۴۹ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۵۰ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۵۱ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۵۲ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۵۳ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۵۴ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۵۵ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۵۶ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۵۷ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۵۸ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۵۹ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۶۰ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۶۱ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۶۲ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۶۳ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۶۴ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۶۵ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۶۶ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۶۷ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۶۸ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۶۹ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۷۰ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۷۱ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۷۲ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۷۳ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۷۴ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |
| | | ۷۵ | ۲۵ | ۲۵ | ۲۳ | ۴۹ | ۵۰ |

با توجه به نتایج جدول ۳، از شش مقاله مورد جست‌وجو در پایگاه «وب‌آوساینس»، ۷۵ تصویر به روش بصری از مقاله‌ها استخراج شد. با استفاده از نرم‌افزار Extreme Picture Finder، معلوم شده که از ۱۲۲ تصویر بازیابی شده توسط نرم‌افزار، ۸۴ مورد با عنوان‌های مقاله‌ها مرتبط بودند. همچنین، جهت بررسی میزان شباهت بین تصویرهای موجود در مقاله‌ها و تصویرهای بازیابی شده توسط نرم‌افزار از حداقل شباهت ۵۰ درصدی استفاده شده است و در کل، میانگین درصد شباهت بین تصویرهای موجود در مقاله‌ها و تصویرهای بازیابی شده توسط نرم‌افزار ۵۸/۲۵ است. از تعداد ۲۳۴ تصویر بازیابی شده از مقاله‌ها توسط نرم‌افزار، ۱۶۷ مورد با چکیده مقاله‌ها مرتبط بودند. با بررسی میزان شباهت حداقل ۵۰ درصدی بین تصویرهای موجود در مقاله‌ها و تصویرهای بازیابی شده توسط نرم‌افزار ۴۶/۱۶ درصد شباهت وجود دارد. از تعداد ۱۸۸ تصویر بازیابی شده از مقاله‌ها توسط نرم‌افزار، ۸۸ مورد با بیان مسئله مقاله‌ها مرتبط بودند. با بررسی میزان شباهت حداقل ۵۰ درصدی بین

| ردیف | عنوان مقاله‌ها | تعداد تصویرهای موجود در مقاله‌ها | عنوان | تعداد تصویرهای بازیابی شده | چکیده | بیان مسئله | نتیجه گیری |
|------|---|----------------------------------|-------|----------------------------|--------------------------------|----------------------|------------|
| ۱ | Hand Gesture Recognition for Human Computer Interaction | ۷ | ۲۴ | ۲۴ | تعداد و درصد شباهت بین تصویرها | تعداد تصویرهای مرتبط | ۰ |
| ۲ | Research and Implementation of computer Date Security Management System | ۲ | ۲۸ | ۲۸ | تعداد و درصد شباهت بین تصویرها | تعداد تصویرهای مرتبط | ۰ |
| ۳ | Analysis of Computer Fashion Illustration's Prospects | ۲ | ۲۵ | ۲۵ | تعداد و درصد شباهت بین تصویرها | تعداد تصویرهای مرتبط | ۰ |
| ۴ | Graphic Design on Education Computer Games | ۴ | ۲۱ | ۲۵ | تعداد و درصد شباهت بین تصویرها | تعداد تصویرهای مرتبط | ۰ |
| جمع | | ۷۵ | ۱۲۳ | ۱۸۸ | | | ۳۶ |
| | میانگین (درصد) | ۵۸/۲۵ | ۴۹ | ۴۶/۱۶ | | | ۳۹/۰۸ |

با توجه به نتایج جدول ۴، از شش مقاله مورد جست‌وجو در پایگاه «اسکوپوس»، ۳۰ تصویر به روش بصری از مقاله‌ها استخراج شد. با استفاده از نرم‌افزار Extreme Picture Finder، از ۱۵۷ تصویر بازیابی‌شده توسط نرم‌افزار، ۱۲۲ مورد با عنوان‌های مقاله‌ها مرتبط بودند. همچنین، جهت بررسی میزان شباهت بین تصویرهای موجود در مقاله‌ها و تصویرهای بازیابی‌شده توسط نرم‌افزار از حداقل شباهت ۵۰ درصدی استفاده شده است و

در کل، میانگین شباهت بین تصویرهای موجود در مقاله‌ها و تصویرهای بازیابی شده توسط نرم‌افزار ۱۸/۸۳ درصد است.

از تعداد ۱۵۵ تصویر بازیابی شده از مقاله‌ها توسط نرم‌افزار، ۹۳ مورد با چکیده مقاله‌ها مرتبط بودند. با بررسی میزان شباهت حداقل ۵۰ درصدی بین تصویرهای موجود در مقاله‌ها و تصویرهای بازیابی شده توسط نرم‌افزار ۱۹/۷۵ درصد شباهت وجود دارد. از تعداد ۱۳۶ تصویر بازیابی شده از مقاله‌ها توسط نرم‌افزار، ۸۵ مورد با بیان مسئله مقاله‌ها مرتبط بودند. با بررسی میزان شباهت حداقل ۵۰ درصدی بین تصویرهای موجود در مقاله‌ها و تصویرهای بازیابی شده توسط نرم‌افزار ۲۶/۰۸ درصد شباهت وجود دارد. از تعداد ۱۵۲ تصویر بازیابی شده از مقاله‌ها توسط نرم‌افزار، ۱۰۷ مورد با نتیجه‌گیری مقاله‌ها مرتبط بودند. با بررسی میزان شباهت حداقل ۵۰ درصدی بین تصویرهای موجود در مقاله‌ها و تصویرهای بازیابی شده توسط نرم‌افزار ۹/۵۸ درصد شباهت وجود دارد. در نتیجه، در پایگاه اطلاعاتی «اسکوپوس» بیشترین درصد شباهت تصویرهای بازیابی شده توسط نرم‌افزار با قسمت بیان مسئله مقاله‌ها و کمترین درصد شباهت با قسمت نتیجه‌گیری بوده است.

برای پاسخگویی به پرسش میزان مطابقت محتوا با تصویرها در پایگاه‌های اطلاعاتی علمی (عنوان، چکیده، بیان مسئله و نتیجه‌گیری) مقاله‌ها به صورت جداگانه وارد نرم‌افزار Extreme Picture Finder گردیده و تصویرها استخراج و مورد تجزیه تحلیل قرار گرفتند. یافته‌ها نشان می‌دهد که بیشترین میزان مطابقت محتوا با تصویرها به ترتیب، در پایگاه‌های «وب‌آوساینس» ۴۸/۱۲ درصد، «مدلاین» ۲۶/۵۱ درصد، «اسکوپوس» ۱۸/۵۶ درصد و «ساینس دایرکت» ۱۴/۲ درصد است.

با توجه به یافته‌های پژوهش، میانگین میزان مشابهت عنوان با تصویرها در پایگاه «مدلاین» ۴۱/۰۸ (جدول ۱)، در پایگاه «ساینس دایرکت» ۱۱/۱۶ (جدول ۲)، در پایگاه «وب‌آوساینس» ۵۸/۲۵ (جدول ۳)، و در پایگاه «اسکوپوس» ۱۸/۸۳ (جدول ۴) است. طبق این یافته‌ها، تصویرها با عنوان‌های مقاله‌ها در پایگاه «وب‌آوساینس» بیشترین میزان مطابقت و در پایگاه «ساینس دایرکت» کمترین میزان مطابقت را دارند.

با توجه به یافته‌های پژوهش، میانگین درصد میزان مشابهت چکیده با تصویرها در پایگاه «مدلاین» ۲۵/۳۳ درصد (جدول ۱)، در پایگاه «ساینس دایرکت» هیچ تشابه وجود نداشت (جدول ۲)، در پایگاه «وب‌آوساینس» ۴۶/۱۶ درصد (جدول ۳) و در پایگاه «اسکوپوس» ۱۹/۷۵ درصد (جدول ۴) است. طبق این یافته‌ها تصویرها با چکیده مقاله‌ها

در پایگاه «وب آوساینس» بیشترین میزان مطابقت و در پایگاه «ساینس دایرکت» کمترین میزان مطابقت را دارند.

با توجه به یافته‌های پژوهش، میانگین درصد میزان مشابهت بیان مسئله با تصویرها در پایگاه «مدلاین» ۱۸/۳۳ درصد است (جدول ۱)، در پایگاه «ساینس دایرکت» هیچ شباهتی وجود نداشت (جدول ۲)، در پایگاه «وب آوساینس» ۴۹ درصد (جدول ۳) و در پایگاه «اسکوپوس» ۲۶/۰۸ درصد (جدول ۴) است. بر اساس این یافته‌ها، تصویرها با بیان مسئله مقاله‌ها در پایگاه «وب آوساینس» بیشترین میزان مطابقت و در پایگاه «ساینس دایرکت» کمترین میزان مطابقت را دارد.

با توجه به یافته‌های پژوهش، میانگین درصد میزان مشابهت نتیجه‌گیری با تصویرها در پایگاه «مدلاین» ۲۱/۴۲ (جدول ۱)، در پایگاه «ساینس دایرکت» ۱۷/۲۵ (جدول ۲)، در پایگاه «وب آوساینس» ۳۹/۰۸ (جدول ۳) و در «پایگاه اسکوپوس» ۱۸/۸۳ (جدول ۴) است. بر اساس یافته‌های این پژوهش، پایگاه اطلاعاتی «وب آوساینس» بیشترین میزان شباهت و پایگاه «اسکوپوس» کمترین میزان شباهت میان تصویرها و نتایج مقاله‌ها را داشتند.

هدف دیگر در این پژوهش شناسایی چالش‌های بازیابی و تطبیق محتوا با تصویرها در پایگاه‌های مورد بررسی بود. با بررسی نتایج خروجی نرم‌افزارها و بررسی دستی جهت شناسایی چالش‌ها، آنچه به دست آمد، مشکل اصلی فقدان استانداردهای مصورسازی در به‌کارگیری تصویرها در مقاله‌های پایگاه‌ها بود. چالش دیگر، نبود قابلیت‌های جامع نرم‌افزارهای مورد استفاده است؛ هر چند این نرم‌افزارها از میان ۴۰ نرم‌افزار با دقت انتخاب شده بودند. مشکل دیگر فقدان شیوه‌نامه استاندارد در چیدمان تصویرها در مقاله‌هاست. در این بررسی در بیشتر مقاله‌ها به وضوح دیده شد که بین متن و تصویرهای به‌کاررفته سنخیت محتوایی پایینی حاکم است.

۵. نتیجه‌گیری و بحث

هدف اصلی این پژوهش بررسی مطابقت محتوا با تصویرها در پایگاه‌های اطلاعاتی علمی (مدلاین، ساینس دایرکت، وب آوساینس و اسکوپوس) است. این پژوهش در صدد بود تا بتواند میزان مطابقت محتوای مقاله‌ها با تصویرها را در پایگاه‌های اطلاعاتی علمی بررسی کند و نشان دهد که میان محتوا با تصویرها در پایگاه‌های مذکور به چه میزان ارتباط وجود دارد. در این پژوهش، مقاله‌ها و تصویرها داده‌های ورودی، و داده‌های

خروجی تصویرهای مرتبط با محتوای مقاله‌های اصلی در پایگاه‌های اطلاعاتی هستند. با انجام این پژوهش مشخص شد که به ترتیب، بیشترین میزان شباهت محتوا با تصویرها در پایگاه «وب آوساینس» (۴۸/۱۲ درصد)، پایگاه «مدلاین» (۲۶/۵۱ درصد)، پایگاه «اسکوپوس» (۱۸/۵۶ درصد) و پایگاه «ساینس دایرکت» (۱۴/۲ درصد) وجود دارد.

دلیل احتمالی بالاترین درصد شباهت محتوا با تصویرها در پایگاه «وب آوساینس»، گستره پوشش موضوعی این پایگاه است که از تنوع موضوعی مقاله‌ها در حوزه‌های مختلف برخوردار است. این پایگاه در هر چهار بخش عنوان، چکیده، بیان مسئله و نتیجه‌گیری مقاله‌ها، از درصد شباهت بیشتری نسبت به سایر پایگاه‌ها برخوردار است. پایگاه «مدلاین» هم به دلایل پوشش موضوعی پزشکی و کثرت تصویرهای پزشکی در مقاله‌ها در جایگاه بعدی قرار دارد.

همچنین، میزان درصد‌های ارائه‌شده در قسمت بحث نشان می‌دهد که به ترتیب، بین عنوان‌ها و تصویرهای مقاله‌ها در پایگاه‌ها (۳۳/ درصد ۳۲)، بیان مسئله و تصویرهای مقاله‌ها در پایگاه‌ها (۲۳/۳۵ درصد)، چکیده و تصویرهای مقاله‌ها در پایگاه‌ها (۲۲/۷۵ درصد) و نتیجه‌گیری و تصویرهای مقاله‌ها در پایگاه‌ها (۲۱/۸۳ درصد) بیشترین شباهت وجود داشته است. دلیل احتمالی تشابه بین عنوان‌ها و تصویرهای مقاله‌ها را می‌توان وجود کلیدواژه‌های اصلی در عنوان بیان کرد.

نتایج پژوهش حاضر با پژوهش (Unar, Wang & Zhang (2018) از جهت بازیابی بر اساس ویژگی‌های بصری و متنی از تصویر و ارائه یک روش جدید جست‌وجوی تصویر با تأکید بر کلمات کلیدی و ترکیبی از آن‌ها و با پژوهش (Jenni, Mandala and Sunar (2015) که به طبقه‌بندی پایگاه داده و انتخاب ویژگی کدگذاری رشته رنگی در بازیابی تصویر پرداخته است، همچنین با پژوهش «بیات» و همکاران (۱۳۹۳) که استفاده هم‌زمان از ویژگی بافت و رنگ تصویر باعث کارایی بالای نظام بازیابی اطلاعات است، همسو و مرتبط است. اما پژوهش حاضر به مرحله پس از بازیابی و بازخورد میزان شباهت تصویرهای بازیابی شده با محتوا می‌پردازد، در حالی که رویکرد پژوهش‌های (Unar, Wang and Zhan (2018) و (Jenni, Mandala and Sunar (2015) و «بیات و همکاران» (۱۳۹۳) به فرایند بازیابی است و به مرحله پس از بازیابی نمی‌پردازند. این پژوهش با پژوهش «آزودینا و هاجو» که راهکار بازخورد ارتباطی تصویرها جهت بهبود دقت سیستم‌های بازیابی مبتنی بر محتوا را مورد بررسی قرار می‌دهد (Azodinia and Hajdu 2016)، و با پژوهش «مشرم، تاکار و گداده» که

تأثیر الگوریتم ترکیبی (ای‌سی‌پی‌اس‌ا) در بازیابی تصویرهای مبتنی بر محتوا را اثبات کرد (Meshram, Thakare, and Gudadhe 2016) و پژوهش «کا» که روش جدید بازیابی تصویر مبتنی بر محتوا بر اساس نظریه اطلاعات را معرفی کرد (Gao et al. 2014)، و همچنین، با پژوهش «شمسی گوشکی» و همکاران که روش بازخورد ربط بر اساس بهبود تابع شباهت در بازیابی تصویر بر اساس محتوا را روش برتر معرفی نمود (۱۳۹۳)، ارتباط نزدیکی دارد. اما در بررسی سنجش میزان شباهت بین تصویر و محتوا نتایج متفاوتی به دست آمده است. نتایج پژوهش حاضر با پژوهش «ذوالفقاری و خسروی» که به ارتقای بازیابی تصویر و دقت در بازیابی اشاره دارد (۱۳۹۲) و با پژوهش «حیاتی، سریزدی و نظام‌آبادی‌پور» که به جداسازی تصویر به مؤلفه‌های بافت و ساختار برای بازیابی تصویر بر اساس محتوا پرداخته (۱۳۹۱)، کاملاً ارتباط دارد.

با توجه به شاخص‌های فنی نرم‌افزارهای مورد استفاده در این پژوهش، از دیدگاه فنی، استانداردسازی مصورسازی تصویرهای به کاررفته در مقاله‌های پایگاه‌های علمی می‌تواند مؤثرترین روش در بازیابی محتوا با تصویرها با میزان تشابه بالا باشد. از آنچه در حین بازیابی محتوا و تصویرها و مطابقت بین آن‌ها انجام شد، به وضوح مشخص شد تصویرهایی که از استانداردهای مصورسازی (متن، رنگ، لبه، حاشیه و ...) به دور بودند، بازیابی نشده و در صورت بازیابی از میزان شباهت نازلی برخوردار بودند. هر چند در این فرایند نقاط ضعفی هم متوجه نرم‌افزارهای مورد استفاده است، اما عمده مشکل به ساختار فنی تصویرهای به کاررفته مربوط است.

علاوه بر مشکلات فوق‌الذکر، مشکل دیگری که در انجام این پژوهش مشخص شد، چیدمان تصویرها در مقاله‌هاست. برخی تصویرها بدون توجه به متن در انتهای مقاله آورده شده‌اند. این مورد در پایگاه «وب‌آوساینس» کمتر دیده می‌شود، اما در سایر پایگاه‌های مورد بررسی به کثرت دیده شد. در بازبینی دستی جهت حصول اطمینان از نتایج نرم‌افزار که در برخی موارد به صورت تصادفی انجام شد، مشکلات درج تصویرها در متن و نبود شیوه‌نامه‌ای مشخص در این خصوص محرز شد. این پژوهش پیشنهادی کاربردی زیر را ضروری می‌داند:

۱. طراحی یک نرم‌افزار جامع جهت انجام تطابق و ایجاد شباهت محتوا و تصویرهای مقاله‌های پایگاه‌ها؛
۲. استانداردسازی تصویرها و درج مقاله‌های پایگاه‌های اطلاعاتی؛
۳. توجه به مصورسازی اطلاعات و اهمیت آن در درک بهتر محتوا.

- همچنین، پژوهش‌هایی برای آینده به شرح زیر پیشنهاد می‌شود:
- ◇ انجام پژوهشی مشابه بر روی سایر پایگاه‌های اطلاعات علمی و مقایسه نتایج حاصل با نتایج این پژوهش؛
 - ◇ انجام پژوهش بر روی پایگاه‌های اطلاعات علمی این پژوهش با مقالات متفاوت؛
 - ◇ استفاده از نرم‌افزارهای دیگر در بازیابی و بررسی میزان شباهت بین تصاویر بازیابی شده و محتوا.
 - ◇ این پژوهش با محدودیت‌هایی همراه بود که بارزترین آن‌ها عبارت‌اند از:
 - ◇ عدم دسترسی به نرم‌افزارهای جامع و کاملاً استاندارد بازیابی و تطبیق شباهت به دلیل مسایل مالی و محدودیت دسترسی؛
 - ◇ زمان‌بر بودن تطبیق دستی بین تصویرهای بازیابی شده و محتوا جهت حصول اطمینان از نتایج تطبیق ماشینی؛
 - ◇ عدم یکدستی تصویرهای به کاررفته در مقالات از یک روش استاندارد مشخص؛
 - ◇ عدم مصورسازی استاندارد تصویرها.

فهرست منابع

- بیات، اکرم، حسن ختن‌لو، امیرحسین براتی، و محسن شاکری. ۱۳۹۳. بازیابی مبتنی بر محتوای تصویرها با استخراج ویژگی‌های بافت و رنگ تصویر با استفاده از مورفولوژی و هیستوگرام رنگی. دومین کنفرانس بین‌المللی بازشناسی الگو و تحلیل تصویر. رشت، انجمن ماشین‌بینایی و پردازش تصویر ایران، دانشگاه گیلان.
- پوررضا، علیرضا. ۱۳۹۴. بازیابی تصویر بر اساس محتوای داده‌های کلان. پایان‌نامه کارشناسی ارشد، دانشگاه سمنان.
- حیاتی، سعید، سعید سریزدی، و حسین نظام‌آبادی‌پور. ۱۳۹۱. جداسازی تصویر به مؤلفه‌های بافت و ساختار برای بازیابی تصویر بر اساس محتوا. نشریه مهندسی برق و مهندسی کامپیوتر ایران ۱۰ (۲): ۱۱۵-۱۲۵.
- ذوالفقاری، احمد، و حسین خسروی. ۱۳۹۲. روشی سریع در بازیابی تصویرها مبتنی بر محتوا با استفاده از ترکیب ویژگی لبه و رنگ. هشتمین سمپوزیوم مهندسی برق و توسعه پایدار. مشهد. مؤسسه آموزش عالی خاوران.
- شمسی گوشکی، اسما، حسین نظام‌آبادی‌پور، سعید سریزدی، و احسان‌الله کبیر. ۱۳۹۳. روشی برای بازخورد ربط بر اساس بهبود تابع شباهت در بازیابی تصویر بر اساس محتوا. پردازش علائم و داده‌ها ۱۱ (۲): ۴۳-۵۵.

References

- Anuja, kh., and Siddarth Ladhake. 2015. Advanced Image Retrieval with Topical Classification Strategy. *Procedia Computer Science* 48 (1): 263-268.
- Aujol, Jean. F., G. Gilboa, T., Chan, and S. Osher. 2006. Structure-texture image decomposition modeling, algorithms, and parameter selection. *International journal of computer vision* 67 (1): 111-136.
- Azodinia, Mohammadreza, and Andras Hajdu. 2016. A Novel Combinational Relevance Feedback Based Method for Content-based Image Retrieval. *Acta Polytechnica Hungarica* 5 (13): 121-134.
- Cao, Wenming., N. Liu, Q. Kong, and H. Feng. 2014. Content-based image retrieval using high-dimensional information geometry. *Science China Information Sciences* 7 (57): 1-17.
- Hassan, Ibrahim, and jin Zhang. 2001. Image search engine feature analysis. *Online Information Review* 25 (2): 103-114.
- Goodrum, Abbya A. 2000. Image information retrieval: An overview of current research. *Informing Science* 3 (2): 63-66.
- Jenni, Kommineni., S. Mandala, and Mohd S. Sunar. 2015. Content Based Image Retrieval using colour strings comparison. *Procedia Computer Science* 50 (1): 374-379.
- Jou, Chichang., and Hung C. Lee. 2004. Handwritten numeral recognition based on simplified feature extraction, structural classification and fuzzy memberships. In *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*. Springer, Berlin, Heidelberg, pp.372-381.
- Kherfi, Mohammed L., Djemel Ziou, and Alan Bernardi. 2004. Image retrieval from the world wide web: Issues, techniques, and systems. *ACM Computing Survey.(CSUR)* 36 (1): 35-67.
- khodaskar, A. & S. Ladhake. 2015. Advanced image retrieval with topical classification strategy. *Procedia Computer Science* 48: 263-268.
- Li, Jing., and Nigel M., Allinson. 2013. Relevance Feedback in Content-Based Image Retrieval: A Survey, *Handbook on Neural Information Processing*, Springer Berlin Heidelberg, pp. 433-469.
- Liu, Ying, D. Zhang, G. Lu, and W. Y Ma. 2007. A survey of content-based image retrieval with high-level semantics. *Pattern recognition* 40 (1): 262-282.
- Meshram, Shubhangi P., Anuradha D. Thakare, and Santwana Gudadhe. 2016. Hybrid swarm intelligence method for post clustering content based image retrieval. *Procedia Computer Science* 79 (1): 509-515.
- Nezamabadi-pour, Hossein, and Ehsanollah Kabir. 2004. Image retrieval using histograms of uni-color and bi-color blocks and directional changes in intensity gradient. *Pattern Recognition Letters* 25 (14): 1547-1557.
- Smeulders, Arnold W. M., M. Worring, S. Santini, A. Gupta, and R. Jain. 2000. Content-based image retrieval at the end of the early years. *IEEE Transactions on pattern analysis and machine intelligence* 22 (12): 1349-1380.
- Unar, Salahuddin, Xingyuan Wang, and Chuan Zhang. 2018. Visual and textual information fusion using Kernel method for content based image retrieval. *Information Fusion* 44: 176-187.

عادل سلیمانی نژاد

دارای مدرک دکتری علم اطلاعات و دانش‌شناسی است. ایشان هم‌اکنون استادیار بخش علم اطلاعات و دانش‌شناسی دانشگاه شهید باهنر کرمان است.

داده‌کاوی، متن‌کاوی، نظام‌های ذخیره و بازیابی اطلاعات، و علم‌سنجی از جمله علایق پژوهشی ایشان است.



فریبرز درودی

دارای مدرک دکتری علم اطلاعات و دانش‌شناسی است. ایشان هم‌اکنون استادیار پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک) است.

سواد رسانه‌ای، مصورسازی اطلاعات و علم‌سنجی از جمله علایق پژوهشی ایشان است.



زینب مرتضی پور

دارای مدرک کارشناسی ارشد علم اطلاعات و دانش‌شناسی از دانشگاه شهید باهنر کرمان است.

بازیابی اطلاعات، داده‌کاوی و مدیریت دانش از جمله علایق پژوهشی ایشان است.

