

Comparison of Sakhr and Google Optical Characters for Persian Language

Mehrnaz Khorasanchi

PhD in Knowledge & Information Science; Researcher of Platform Thinking in Libraries and Information Centers; Tehran, Iran;
Email: mkhorasanchi@gmail.com

Tayebeh Shahmirzadi*

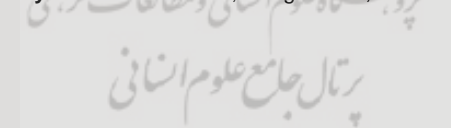
PhD in Knowledge & Information Science; Assistant Professor; Faculty of Agricultural Center for Information Technology and Services; Agricultural Research, Education and Extension Organization (AREEO); Tehran, Iran t.shahmirzadi@areeo.ac.ir

Iranian Journal of
**Information
Processing and
Management**

Received: 01, Sep. 2017 | Accepted: 06, Jan. 2021

Abstract: Concerning the importance & necessity of using OCR programs in library and information centers, this research aims at evaluating the new Persian OCR programs (Sakhr and Google). Analytical exploratory research method is used and data are gathered in a practical way through purposive observing the cases. Checklist is used as the gathering tools. Results of the study show some good and bad specifications in both programs. The outcome part of the Google program is acceptable with fewer errors but in other outcome module characteristics such as word recognition, Sakhr program did better. Due to the availability of "Google" software only through the Internet and the possibility of uploading only 20 pages at a time, there are usage restrictions for the user.

Keywords: Sakhr OCR, Google OCR, Farsi OCR, Document, Evaluation



Iranian Research Institute
for Information Science and Technology
(IranDoc)

ISSN 2251-8223

eISSN 2251-8231

Indexed by SCOPUS, ISC, & LISTA

Vol. 36 | No. 4 | pp. 1181-1202

Summer 2021



* Corresponding Author

مقایسه نویسه خوان‌های نوری صخر و گوگل برای زبان فارسی

مهروز خراسانچی

دکتری علم اطلاعات و دانش‌شناسی؛ پژوهشگر حوزه
تفکر پلتفرمی در کتابخانه‌ها و مراکز اطلاع‌رسانی؛
تهران، ایران mkhorasanchi@gmail.com

طیبه شه میرزادی

دکتری علم اطلاعات و دانش‌شناسی؛ استادیار؛
مرکز فناوری اطلاعات و اطلاع‌رسانی کشاورزی؛
سازمان تحقیقات، آموزش و ترویج کشاورزی؛
تهران، ایران؛
t.shahmirzadi@areeo.ac.ir پدیدآور رابط



دریافت: ۱۳۹۹/۰۶/۱۰ | پذیرش: ۱۳۹۹/۱۰/۱۷ | مقاله برای اصلاح به مدت ۲۰ روز نزد پدیدآوران بوده است.

چکیده: با توجه به اهمیت و ضرورت استفاده از نویسه‌خوان‌های نوری در کتابخانه‌ها و مراکز اطلاع‌رسانی، هدف اصلی پژوهش حاضر ارزیابی نرم‌افزارهای نویسه‌خوان نوری فارسی «صخر» و «گوگل» است. پژوهش انجام‌شده از نوع کاربردی و ابزار گردآوری داده‌ها سیاهه و آرسی است. برای گردآوری معیارها و سیاهه و آرسی پژوهش از روش کتابخانه‌ای و برای انجام پژوهش از شیوه پیمایشی-اکتشافی از طریق مشاهده هدفمند نتایج در نرم‌افزارهای مورد بررسی استفاده شده است. نتایج حاصل از پژوهش نشان داد که هر یک از نرم‌افزارهای مورد بررسی دارای معایب و مزایایی است. در خروجی‌های مربوط به متن، نرم‌افزار «گوگل» با درصد خطای کمتر، مناسب‌تر است؛ ولی در سایر موارد تشخیصی و خروجی، نرم‌افزار «صخر» بهتر عمل می‌کند. به دلیل دسترسی به نرم‌افزار «گوگل» تنها از طریق اینترنت و امکان آپلود هر بار تنها ۲۰ صفحه، محدودیت‌هایی در استفاده برای کاربر وجود دارد.

کلیدواژه‌ها: نویسه‌خوان نوری صخر، نویسه‌خوان نوری گوگل، نویسه‌خوان نوری فارسی، سند، ارزیابی

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

شاپا (جایی) ۲۲۵۱-۸۲۲۳

شاپا (الکترونیکی) ۸۲۳۱-۲۲۵۱

نمایه در SCOPUS، LISTA، و ISI

jipm.irandoc.ac.ir

دوره ۳۶ | شماره ۴ | صص ۱۱۸۱-۱۲۰۲
تابستان ۱۴۰۰



۱. مقدمه و بیان مسئله

واژه «اُسی آر»^۱ یا «نویسه‌خوان نوری» سال‌هاست که در سراسر جهان مورد توجه بسیاری از پژوهشگران قرار گرفته و مقاله‌های مربوط به این موضوع همه‌ساله در بسیاری از کنفرانس‌های معتبری که در زمینهٔ بنیادی ماشین‌آ، شبکه‌های عصبی، هوش مصنوعی، پردازش تصویر و ... تشکیل می‌شود، در سطح جهان ارائه می‌گردد. بسیاری از مسائل مطرح‌شده در حوزهٔ نویسه‌خوانی نوری را می‌توان با حوزهٔ نوشتاری زبان (و حتی دیگر حوزه‌های زبان) مرتبط دانست. هرچند در کشور ما و برخی کشورهای دیگر بر روی نویسه‌خوان‌های نوری فارسی کارهای ارزشمندی انجام شده، اما می‌توان گفت زبان‌های لاتین، چینی و ژاپنی بسیار پیش از زبان فارسی به ایده‌آل نزدیک شده‌اند (دبیرخانهٔ شورای عالی اطلاع‌رسانی ۱۳۸۸).

در چند دههٔ گذشته، بازشناسی الگوهای نوشتاری شامل حروف، ارقام و دیگر نمادهای متداول در اسناد نوشته‌شده به زبان‌های مختلف، توسط گروه‌های مختلفی از پژوهشگران مورد مطالعه و بررسی قرار گرفته است. نتیجهٔ این پژوهش‌ها منجر به ایجاد مجموعه‌ای از روش‌های سریع و تا حد زیادی مطمئن موسوم به «اُسی آر» یا «بازشناسی نوری حروف» به‌منظور رایانه‌ای کردن اطلاعات موجود در اسناد، مدارک، کتاب‌ها و سایر نوشتارهای حروف چینی و حتی دست‌نوشته شده است. اصطلاح «اُسی آر» به فوننی اطلاق می‌شود که در تصاویر اسکن یا فکس‌شده نواحی متنی را تشخیص می‌دهند و سپس، این نواحی (تصویری) را به متن قابل ویرایش تبدیل می‌نمایند (Trier 1996 نقل در فرامرزی ۱۳۸۴).

واژه «اُسی آر» در آغاز در مورد بازشناسی ارقام و حروف چاپی به کار گرفته می‌شد، ولی امروزه، بیشتر برای بازشناسی مستندات چاپی مثل صفحه‌های کتاب‌ها، مجله‌ها و نامه‌های چاپی به کار می‌رود. تصویر سند اغلب توسط اسکنر و یا دوربین دیجیتال تولید می‌شود و شامل تعدادی پیکسل با رنگ‌های مختلف و سطوح روشنایی گوناگون است. از دید انسان، یک سند ممکن است ارزش اطلاعاتی زیادی داشته باشد، لیکن از دید رایانه، تصویر یک سند با تصویر یک منظره تفاوتی ندارد؛ چرا که هر دوی آن‌ها مجموعه‌ای از پیکسل‌ها هستند. برای این که بتوان از اطلاعات نوشتاری تصویر سند استفاده کرد باید

1. OCR (Optical Character Recognition)

2. Computer vision

به‌نحوی نوشته‌های موجود در سند را بازشناسی کرد. چنین کاری توسط نرم‌افزارهای نویسه‌خوان نوری انجام می‌شود (دبیرخانه شورای عالی اطلاع‌رسانی ۱۳۸۸). به‌عبارت دیگر، هر فایل تصویری (اسکن‌شده)، حجم زیادی دارد و جست‌وجوی متنی در آن ممکن نیست. این در حالی است که فایل خروجی سامانه نویسه‌خوان بسیار کم‌حجم و قابل جست‌وجوست. اطلاعات فایل متنی را به‌راحتی می‌توان ویرایش و در جای دیگر استفاده کرد، در حالی که اطلاعات موجود در تصویر سند قابل ویرایش نیست. برای وارد کردن مطالب چاپی و یا دست‌نویس در رایانه دو راه وجود دارد: یکی راه این است که اسناد دوباره به‌وسیله اپراتور تایپ شود و راه دیگر، تصویربرداری از روی اسناد کاغذی به‌وسیله اسکنرها و انتقال تصاویر دیجیتالی تولیدشده به سیستم‌های رایانه‌ای است. برای انجام این مرحله از نویسه‌خوان‌های نوری استفاده می‌شود.

یکی از مهم‌ترین نرم‌افزارهای تجاری نویسه‌خوان نوری که زبان فارسی را پشتیبانی می‌کند، نرم‌افزار «صخر»^۱ است. «صخر» نام یک شرکت مصری است که در زمینه فناوری‌های پردازش زبان طبیعی با تأکید بر زبان و خط عربی کار می‌کند. این نرم‌افزار با هدف پشتیبانی تخصصی از خط و زبان عربی ساخته شده و افزون بر زبان عربی، از برخی زبان‌های دیگر که با خط عربی یا مشتقات آن نوشته می‌شوند (فارسی، اردو) نیز پشتیبانی می‌کند. این نرم‌افزار موفقیت‌های چشمگیری در عرصه فناوری پردازش زبان عربی به‌دست آورده است. محصولات این شرکت طیف وسیعی از ترجمه ماشینی، تولید صوت ماشینی، تبدیل تصویر به متن، و کاربردهایی همچون تحلیل مستندات گفتار، موتور جست‌وجوی وب، سیستم‌های مدیریت اطلاعات در وب و موارد دیگر را دربرمی‌گیرد. در این راستا، شرکت «گوگل» نیز به‌تازگی اقدام به اضافه کردن زبان فارسی در بخش «اُسی‌آر» خدمات تحت وب «گوگل‌درایو» کرده است.^۲ «گوگل» در حال حاضر در زمینه تشخیص نوری حروف به موفقیت چشم‌گیری دست یافته و روی بیش از ۲۰۰ زبان مختلف کار کرده است. بدین ترتیب، ویرایش و جست‌وجو در متون زبان‌های مختلف امکان‌پذیر می‌شود. شناسایی ۲۵ سیستم نوشتاری مختلف هم از مزایای فناوری «اُسی‌آر» این شرکت است.

1. <http://www.sakhr.com/index.php/en/>

2. <https://chrome.google.com/webstore/detail/online-ocr/gcholkakdmcflfbbkdbcocnfggfmcc>

در حوزه خط و زبان فارسی در محیط رایانه‌ای فراهم آوردن امکان ذخیره و بازیابی اطلاعات به خط و زبان فارسی و بهینه‌سازی آن از اولویت فراوانی برخوردار است. نویسه‌خوان نوری به‌عنوان تنها ابزار بازیابی اطلاعات متنی از تصویر، یکی از مهم‌ترین ابزارهای تبدیل اطلاعات موجود به‌صورت قابل استفاده و پردازش در رایانه‌ها و در نتیجه، یکی از ارکان مهم پژوهش محیط رایانه‌ای فارسی به‌شمار می‌آید.

تعداد زیادی از نرم‌افزارهای نویسه‌خوان نوری برای پشتیبانی از زبان فارسی ساخته و منتشر شده‌اند، ولی شوربختانه، اکثر آن‌ها جنبه تبلیغاتی داشته و در عمل، بسیار ضعیف و پر از خطا هستند. با وجود پژوهش‌های فراوان و سرمایه‌گذاری‌های بسیار در این زمینه، پیچیدگی ذاتی مسائل مطرح در این حوزه و گستردگی علوم و فناوری‌های مرتبط با آن سبب شده است که راه زیادی تا تولید نویسه‌خوان‌های نوری ایده‌آل باقی باشد.

پژوهش حاضر با تهیه سیاهه واری برای ارزیابی دو نرم‌افزار نویسه‌خوان «صخر» و «گوگل»، ضمن فراهم‌ساختن امکان انتخاب و استفاده از «آسی‌آر» مناسب‌تر برای پژوهشگران، به تولیدکنندگان این نرم‌افزارها و سایر نرم‌افزارهای مشابه نیز این امکان را می‌دهد که نقاط ضعف خود را شناسایی و برطرف نمایند.

در این مقاله با توجه به اهمیت و ضرورت استفاده از نویسه‌خوان‌های نوری در کتابخانه‌ها و مراکز اطلاع‌رسانی سعی می‌شود که دو نوع از نویسه‌خوان‌های نوری فارسی «صخر» و «گوگل» مورد ارزیابی قرار گرفته و نتایج ارزیابی برای استفاده کاربردی از آن‌ها در کتابخانه‌ها و مراکز اطلاع‌رسانی به‌کار برده می‌شود.

هدف اصلی پژوهش حاضر ارزیابی نرم‌افزارهای نویسه‌خوان نوری فارسی «صخر» و «گوگل» است. هدف‌های فرعی این پژوهش شناسایی صحت تشخیص حرف و کلمه، میزان خطا، سرعت، تشخیص فونت و بازشناسی تصویر سند در نرم‌افزارهای مورد بررسی است.

۲. پرسش‌های پژوهش

۱. امکان‌ات و قابلیت‌های دقت بازشناسی، صحت تشخیص حرف و کلمه، میزان خطا، و

سرعت در نرم‌افزارهای مورد بررسی تا چه میزان کارایی دارند؟

۲. قابلیت‌های تشخیص فونت، بازشناسی زبان‌های مختلف و قابلیت حفظ قالب‌بندی

سند در نرم‌افزارهای مورد بررسی چگونه است؟

۳. قابلیت‌های پردازش گروهی فایل‌ها، دسترسی و کار با دستگاه‌های پوشگر در نرم‌افزارهای مورد بررسی چگونه است؟
۴. توانایی نرم‌افزارهای مورد بررسی در پیش‌پردازش، جداسازی و تشخیص تصاویر و شناسایی طرح‌بندی سند، و ارائه طرح‌بندی در خروجی چگونه است؟
۵. امکانات و قابلیت‌های انتخاب نوع فرمت، تصحیح املائی، بازشناسی، پیکربندی و تعامل با دستگاه‌های ورودی در نرم‌افزارهای مورد بررسی چگونه است؟

۳. پیشینه پژوهش

در زمینه نرم‌افزارهای نویسه‌خوان نوری پژوهش‌های قابل ملاحظه‌ای انجام شده، اما تحقیقات در مورد ارزیابی این نرم‌افزارها محدود است. در زیر به برخی از آن‌ها اشاره می‌شود: «هنروان» در پژوهشی به مزایای روش «اُسی آر» و کاربردهای آن در حوزه کتابداری و اطلاع‌رسانی اشاره داشته و به معرفی و مقایسه چهار برنامه «اُسی آر» (OmniPage Professional 5.0, Recognita Plus 2.0 International, Xerox Imaging Systems, WordScan Plus) که تحت برنامه ویندوز «مایکروسافت» کار می‌کنند، پرداخته است (۱۳۷۵). «فرامرزی» علاوه بر معرفی ویژگی‌ها و پیچیدگی‌های مختص نگارش زبان فارسی که هر نرم‌افزار «اُسی آر» باید در عملیات پردازشی خود لحاظ نماید، نرم‌افزارهای معروف تجاری «اُسی آر» لاتین و فارسی را معرفی، و قابلیت‌ها و نقاط قوت و ضعف آن‌ها را تشریح نمود (۱۳۸۴). «دبیرخانه شورای عالی اطلاع‌رسانی» در پژوهشی افزون بر بررسی مسائل مطرح در حوزه نویسه‌خوان‌های نوری و نگاه به این مسائل از ابعاد مختلف فنی، مدیریتی و کاربردی، به تعریف تعدادی از مهم‌ترین پروژه‌های زیرساختی در زمینه نویسه‌خوانی نوری فارسی نیز پرداخته است (۱۳۸۸).

«خسروی و کبیر» در مقاله‌ای به ارزیابی روش‌های بازشناسی متون فارسی بر مبنای شکل کلی زیرکلمات پرداختند. دو رویکرد برای بازشناسی متون فارسی با استفاده از شکل کلی زیرکلمات ارائه شده و ضمن مقایسه آن‌ها، مزایا و معایب استفاده از آن‌ها نیز بیان شده است. نتایج حاصل حاکی از آن است که رویکرد با نقطه، دقتی حدود ۹۴ درصد و رویکرد بدون نقطه، دقتی حدود ۹۳ درصد در سطح کلمه ارائه می‌دهد، لیکن

در برخورد با تصاویر کم کیفیت و دارای نویزی دچار افت دقت می‌شوند و این کاهش دقت در برخی موارد بسیار شدید است. دلایل این کاهش دقت مورد ارزیابی قرار گرفته و راهکاری برای بهبود آن ارائه شده است. همچنین، ضمن مقایسه دو رویکرد، مزایا و معایب بازشناسی بر مبنای شکل کلی ارائه شده است (۱۳۸۸).

«زند» در پژوهش خود روشی ترکیبی برای تشخیص حروف فارسی ارائه کرده است که در آن جداسازی حروف (که منبع اصلی خطا در سیستم‌های تشخیص حروف فارسی کنونی است)، به صورت کامل انجام نشده و فقط حرف اول هر زیر کلمه جدا شده و تشخیص داده می‌شود (۱۳۹۱).

«سلیمانی و حسینی» در پژوهش خود یک رویکرد مبتنی بر نویسه (یکی از حروف الفبا) را برای شناسایی و اصلاح نویسه فارسی و عربی آفلاین پیشنهاد می‌کنند. آن‌ها مدل ترکیبی بتا-بیضی را برای روش خود انتخاب نموده و یک کتاب را کد کامل و چهار کتاب را کد جزئی زده و تست کرده‌اند. این پژوهش اولین مطالعه برای درون‌یابی مفهوم کتاب‌های کد ترکیبی مبتنی بر مدل شناسایی و اصلاح نویسه‌های فارسی و عربی است (۱۳۹۴).

«ایرانپور مبارکه و احمدی فرد» در پژوهش خود چند ویژگی مختلف مبتنی بر ساختار نگارش زبان فارسی جهت استفاده در بازیابی کلمات تصویری فارسی معرفی کردند. در این پژوهش نتایج حاصل از روش‌های مبتنی بر ویژگی‌های پیشنهادی مقایسه شده است. در ارائه ویژگی‌ها، ساختار نگارشی زبان فارسی در نظر گرفته شده است تا بهترین مطابقت را با روش نگارش فارسی داشته باشد. آزمایش‌های تجربی روی دو مجموعه داده استاندارد «فارسا» و «ایران شهر» مورد بررسی قرار گرفت (۱۳۹۵).

«آگاهی و صالحی» سیستم‌های بازشناسی ارقام دستنویس را بر اساس معیارهایی همچون مقاومت در برابر نویز، چرخش، نیاز به الگوریتم‌های کاهش ابعاد و ترکیب با ویژگی‌های دیگر مورد بررسی و ارزیابی قرار دادند. و در نهایت، به این نتیجه رسیدند که ترکیب ویژگی‌ها می‌تواند تأثیر مستقیمی در کارایی سیستم داشته باشد (۱۳۹۵).

«کانونگو، مارتون و بولبول» در پژوهشی بیان داشتند که بسیاری از شرکت‌های تولیدکننده نویسه‌خوان نوری ادعا می‌کنند که محصولات آن‌ها از لحاظ وضوح دارای کیفیت بسیار بالا هستند (نزدیک به ۹۹/۹ درصد). البته، این رقم در عمل به سختی به دست

می‌آید. در این پژوهش دو محصول عربی شناخته‌شده «صخر» و «اومنی‌پیج»^۱ مورد بررسی قرار گرفته‌اند. آن‌ها در این ارزیابی به این نتیجه رسیده‌اند که محصولات «صخر» نسبت به «اومنی‌پیج»، ۱۵/۴۷ درصد کمتر دچار خطا می‌شوند. وضوح تصویر برای «صخر» ۹۰/۳۳ درصد و برای «اومنی‌پیج» ۸۶/۸۹ درصد است. ملاک هر دو محصول وضوح نویسه‌خوان‌های عربی بودند. به‌طور کلی، «صخر» در رابطه با بهبود کیفیت پایین وضوح تصاویر عملکرد بهتری دارد. تکنیک‌های تجسمی این امکان را به کاربران می‌دهد که الگوریتم تصاویر را گسترش داده و نتایج کار را تحلیل کنند (Kanungo, Marton and Bulbul 1998). آن‌ها در پژوهشی دیگر به ارزیابی عملکرد دو محصول نویسه‌خوان نوری عربی «اومنی‌پیج» و «صخر» از نظر توصیف عملکرد سیستم‌های نویسه‌خوان‌های نوری پرداختند. نویسه‌خوان‌ها از نظر پیشرفت‌های فنی و پیش‌بینی عملکرد سیستم‌ها مورد بررسی قرار گرفتند (Kanungo, Marton and Bulbul 1999).

«شایگان و چان» در مقاله خود اشاره دارند که در شناخت کاراکترهای فارسی-عربی دست‌نویس، انتخاب و استخراج شاخص‌ها از گام‌های بسیار مهم در سیستم‌های شناخت الگوها محسوب می‌شوند و با استفاده از یک مطالعه جامع روی سیستم‌های شناخت ارقام عربی-فارسی مجموعه‌ای از ویژگی‌های شناخته‌شده به دست آمد. نتایج تجربی این پژوهش که بر اساس یکی از بزرگ‌ترین دست‌نوشته‌های ارقام فارسی «هدی» است، وضوح ۹۵/۷۰ درصد تشخیص ارقام دست‌نویس و برنامه‌نویسی عددی در فارسی را نشان داد و به این ترتیب، گام بلندی به سوی استانداردسازی کارهای پژوهشی برداشته شد (Shayegan and Chan 2012).

«دیمان و سینگ» در پژوهش خود دو نرم‌افزار «جی‌اُسی آر»^۲ و «تسراکت»^۳ را مقایسه کردند. آن‌ها بر اساس پارامترها و متریک‌های (استاندارد) متفاوت، نتایج متفاوتی را نشان می‌دهند. در این تحقیق دو نرم‌افزار متن‌باز ذکر شده، مقایسه و در نهایت، به توصیف نحوه کار هر دو نرم‌افزار پرداخته شده است. همچنین، نرم‌افزارها بر اساس پارامترهای مختلف از جمله میزان دقت، وضوح، فونت، میزان روشنایی و نوع تصاویر آن‌ها مورد بررسی قرار گرفتند (Dhiman and Singh 2013).

«ساجدی» در پژوهش خود گامی به سوی استانداردسازی کارهای پژوهشی در تشخیص

کاراکتر نوری در زبان فارسی برداشته است. وی نخست، یک پایگاه داده استاندارد از نوشته‌ها از جمله رقم جداشده، علائم جداشده، اعداد چندرقمی، رشته‌های عددی و کد پستی ارائه می‌دهد. نتایج این تجربه با سایر پژوهش‌ها در زمینه شناخت نوشته‌های دستی مقایسه شد. نتایج به دست آمده نشان دادند که پیشرفت زیادی در دقت سیستم در مقایسه با موارد استفاده از روش‌های مشابه دیده می‌شود (Sajedi 2016).

با نگاهی به پژوهش‌های صورت گرفته درمی‌یابیم که انجام پژوهش‌ها در حوزه نویسه‌خوان‌ها رو به گسترش است و می‌توان آن‌ها را به سه دسته کلی تقسیم کرد:

۱. برخی از پژوهش‌های انجام شده به مبانی نظری و جنبه‌های کاربردی مبحث بازشناسی نوری به صورت جامع پرداخته و به معرفی نرم‌افزارهای معروف فارسی و غیر فارسی و تشریح قابلیت‌ها و نقاط قوت و ضعف آن‌ها پرداخته‌اند. پژوهش «فرامرزی» (۱۳۸۴) و «دبیرخانه شورای عالی اطلاع‌رسانی» (۱۳۸۸) در این گروه قرار دارند.

۲. برخی دیگر از پژوهش‌ها به بررسی مشکلات موجود در نویسه‌خوان‌ها و یا ویژگی‌های ساختار نگارشی زبان فارسی اشاره دارند. پژوهش «خسروی و کبیر» (۱۳۸۸)، «زند» (۱۳۹۱)، «سلیمانی و حسینی» (۱۳۹۴)، «آگاهی و صالحی» (۱۳۹۵)، «شایگان و چان» (۲۰۱۲) و «ساجدی» (۲۰۱۶) از این دسته‌اند.

۳. دسته سوم، پژوهش‌هایی هستند که با مقایسه نرم‌افزارهای مختلف در این زمینه، به کارگیری نرم‌افزار مناسب را پیشنهاد می‌کنند؛ مانند پژوهش «هنروران» که به معرفی و مقایسه چهار برنامه «آسی آر» با مقایسه عواملی مانند متن تهیه شده توسط چاپگر چرخ خورشیدی، متن تهیه شده توسط چاپگر جوهرافشان لیزری، متن فاکس شده و متن کپی شده پرداخته است (۱۳۷۵). «ایرانپور مبارکه و احمدی فرد» آزمایش‌های تجربی را روی دو مجموعه داده استاندارد «فارسا» و «ایران شهر» مورد بررسی قرار دادند (۱۳۹۵). «کانونگو، مارتون و بولبول» در پژوهش خود دو نویسه‌خوان «صخر» و «اومنی پیچ» را از نظر دقت، حذف و وضوح تصویر مورد بررسی قرار دادند. آن‌ها همچنین، در سال ۱۹۹۹ به بررسی این دو نویسه‌خوان از نظر پیشرفت‌های فنی و پیش‌بینی عملکرد سیستم‌ها پرداختند (Kanungo, Marton and Bulbul 1999). «دیمان و سینق» دو نرم‌افزار «جی‌آسی آر» و «تسراکت» را بر اساس پارامترهای مختلف از جمله میزان دقت، وضوح، فونت، میزان روشنایی و نوع تصاویر آن‌ها مقایسه نمودند (Dhiman and Singh 2013).

در این میان، جای خالی پژوهش‌های موردی و بومی بیشتر به چشم می‌خورد. بررسی پیشینه‌های داخل و خارج کشور نشان می‌دهد که تاکنون پژوهشی در خصوص بررسی نرم‌افزارهای جدید مانند «صخر» و «گوگل» انجام نشده است. در پژوهش حاضر تلاش می‌شود که دو نوع از نویسه‌خوان‌های نوری جدید فارسی «صخر» و «گوگل» مورد ارزیابی قرار گرفته و نتایج ارزیابی برای استفاده کاربردی از آن‌ها در کتابخانه‌ها و مراکز اطلاع‌رسانی پیشنهاد می‌شود.

۴. روش پژوهش

پژوهش انجام شده از نوع کاربردی است. برای گردآوری معیارهای سیاهه ارزیابی پژوهش از روش کتابخانه‌ای و برای انجام پژوهش از شیوه پیمایشی-اکتشافی از طریق مشاهده، تحلیل و ارزیابی هدمند نرم‌افزارهای مورد بررسی استفاده شده است. روش اکتشافی برای زمینه‌هایی مناسب است که پدیده‌های مهم آن و نحوه اندازه‌گیری این پدیده‌ها هنوز به روشنی مشخص نشده‌اند (حریری ۱۳۸۵). در این پژوهش با توجه به نبود سیاهه ارزیابی مناسب، برای ارزیابی نرم‌افزارهای نویسه‌خوان نوری «صخر» و «گوگل» از سیاهه ارزیابی معقق ساخته استفاده شد. با ارسال سیاهه ارزیابی برای دو نفر از متخصصان حوزه علم اطلاعات و دانش‌شناسی و سه نفر از متخصصان رایانه، در دو مرحله تغییراتی در آن اعمال شد. این سیاهه ارزیابی از ۵ بخش اصلی تشکیل شده است که شامل (۱) دقت بازشناسی، صحت تشخیص حرف و کلمه، میزان خطا، سرعت و استفاده از واژه‌پرداز چندزبانه، (۲) قابلیت‌های تشخیص فونت، بازشناسی زبان‌های مختلف و قابلیت حفظ قالب‌بندی سند، (۳) پردازش گروهی فایل‌ها، دسترسی و کار با دستگاه‌های پوشگر، (۴) توانایی‌های پیش‌پردازش، جداسازی و تشخیص تصاویر و شناسایی طرح‌بندی سند و ارائه طرح‌بندی در خروجی، و (۵) امکانات و قابلیت‌های انتخاب، نوع فرمت، تصحیح املائی، بازشناسی، پیکربندی و تعامل با دستگاه‌های ورودی است. برای انتخاب معیارهای ارزیابی سعی شد معیارهایی انتخاب شوند که کمی بوده و قابل اندازه‌گیری و از معیارهای اصلی و ضروری باشند. جامعه مورد مطالعه را نرم‌افزارهای نویسه‌خوان نوری «صخر» و «گوگل» تشکیل می‌دهند.

۵. مراحل انجام کار

- برای انجام این پژوهش مراحل زیر انجام شده است:
۱. انتخاب ۲۰ صفحه مقاله در فرمت ورد (Word) به عنوان نمونه؛
 ۲. ذخیره صفحات انتخاب شده در قالب ۴ فونت مختلف (B Nazanin, B Mitra, B zar, B Lotus)؛
 ۳. تبدیل صفحات با فونت‌های مختلف به قالب jpg؛
 ۴. دریافت خروجی pdf از فایل های jpg؛
 ۵. ارائه فایل pdf به نرم افزارهای «صخر» و «گوگل»؛
 ۶. بررسی و ارزیابی نرم افزارها بر اساس سیاهه ارزیابی.

۶. تجزیه و تحلیل اطلاعات

در این بخش از پژوهش، داده‌های استخراجی از «سیاهه و آرسی» در پاسخ به پرسش‌های پژوهش حاضر مورد بررسی و تجزیه و تحلیل قرار گرفت. همان‌طور که در مراحل انجام کار بیان شد، از هر مقاله ۲۰ صفحه در قالب «ورد» برای نمونه انتخاب شد. دلیل تعداد صفحات از این جهت بود که در نرم‌افزار نویسه‌خوان «گوگل» در هر نوبت بیش از ۲۰ صفحه پذیرفته نمی‌شد و پژوهشگران نیز این محدودیت را اعمال نمودند. در مجموع، ۱۰ عنوان مقاله به صورت تصادفی از بین نشریات مختلف انتخاب و ۲۰ صفحه نخست از هر یک ذخیره شد. پس از آن، صفحات انتخاب شده از هر مقاله در قالب ۴ فونت مختلف (B Nazanin, B Mitra, B zar, B Lotus) ذخیره شد. سپس، این ۸۰۰ صفحه (۸۰۰=۱۰×۲۰×۴) با فونت‌های مختلف به قالب jpg تبدیل شدند تا بتوان خروجی pdf از فایل های jpg دریافت نمود. بعد از طی این مراحل، فایل های pdf تهیه شده به نرم‌افزارهای «صخر» و «گوگل» داده شد و نرم‌افزارها بر اساس سیاهه ارزیابی، توسط پژوهشگران این مقاله بررسی و ارزیابی شدند که نتایج آن‌ها در ادامه آمده است.

- ۱-۶. امکانات و قابلیت‌های دقت بازشناسی، صحت تشخیص حرف و کلمه، میزان خطا و سرعت در نرم‌افزارهای مورد بررسی تا چه میزان کارایی دارند؟
- دقت بازشناسی، صحت تشخیص حرف و کلمه از موارد مهم در ارزیابی نرم‌افزارهای

فارسی است و سرعت انجام «اُسی آر» نیز از مؤلفه‌های مهمی است که در سهولت کار با آن نقش مؤثری دارد. به همین دلیل، این معیارها برای ارزیابی انتخاب شدند. در جدول ۱، بررسی‌های انجام‌شده در نرم‌افزارهای انتخابی آورده شده است.

جدول ۱. امکانات و قابلیت‌های دقت بازشناسی، صحت تشخیص حرف و کلمه، میزان خطا، سرعت و استفاده از واژه‌پرداز چندزبانه در نرم‌افزارهای مورد بررسی

ردیف	امکانات و قابلیت‌های	نرم‌افزار صخر (درصد)	نرم‌افزار گوگل (درصد)
۱	دقت بازشناسی	۸۹/۶	۹۷/۴
۲	صحت تشخیص حرف	۹۶	۹۹/۶
۳	صحت تشخیص کلمه	۸۹/۶	۹۷/۴
۴	میزان خطا	۱۰/۴	۲/۶
۵	سرعت	هر صفحه ۲ ثانیه	قابل اندازه‌گیری نیست

در زمینه سنجش میزان امکانات و قابلیت‌های موجود در نرم‌افزارهای مورد بررسی در این پژوهش، در زمینه صحت تشخیص حرف و کلمه به ازای هر مجموعه مورد آزمایش، نسبت درصد کلماتی که درست تشخیص داده شده‌اند به کل کلمات مورد آزمایش، میزان صحت تشخیص حرف و کلمه و دقت بازشناسی را نشان می‌دهد. بر اساس اطلاعات درج‌شده در جدول ۱، نرم‌افزار «صخر» از نظر دقت بازشناسی ۸۹/۶ درصد است؛ در صورتی که دقت بازشناسی در نرم‌افزار «گوگل» ۹۷/۴ درصد است. از نظر صحت تشخیص حرف، نرم‌افزار «گوگل» با ۹۹/۶ درصد نسبت به نرم‌افزار «صخر» با ۹۶ درصد عملکرد بهتری دارد. همچنین، از نظر صحت تشخیص کلمه نیز نرم‌افزار «گوگل» با ۹۷/۴ درصد نسبت به نرم‌افزار «صخر» با ۸۹/۶ درصد در جایگاه بالاتری قرار دارد.

برای بررسی میزان خطا در نرم‌افزارهای مورد بررسی در هر مجموعه مورد آزمایش، نسبت تعداد کلماتی که غلط تشخیص داده شده‌اند به کل کلمات مورد آزمایش، میزان خطا را نشان می‌دهد. منظور از غلط، کلماتی است که با کلمه ورودی مطابق نباشند. میزان خطا در «صخر» ۱۰/۴ درصد، ولی در نرم‌افزار «گوگل» ۲/۶ درصد است.

برای سنجش سرعت انجام کار در نرم‌افزارهای مورد بررسی، در نرم‌افزار «صخر» سرعت بر مبنای مدت‌زمان صرف‌شده برای «اُسی آر» یک صفحه محاسبه و برابر با ۲ ثانیه برآورد گردید؛ ولی برای نرم‌افزار «گوگل» که تنها از طریق اینترنت قابل دسترسی

است، این سنجه به دلیل تغییر در سرعت و نوع اینترنت مصرفی متغیر است. از این رو، میزان دقیق آن قابل اندازه گیری نیست.

۲-۶. قابلیت‌های تشخیص فونت، بازشناسی زبان‌های مختلف و قابلیت حفظ قالب‌بندی سند در نرم‌افزارهای مورد بررسی چگونه است؟

اگر سیستم «اُسی آر» از قلم‌های خاصی پشتیبانی کند، با گذر زمان دچار مشکل خواهد شد. این نکته به ویژه برای زبان فارسی که حروف به هم پیوسته زیادی دارد و قلم‌های مختلف آن تفاوت‌های ظاهری زیادی دارند، ضروری است. به همین دلیل، میزان بازشناسی فونت‌های مختلف در نرم‌افزارهای «صخر» و «گوگل» مورد بررسی قرار گرفت و نتایج در جدول ۲، درج شده است.

جدول ۲. قابلیت‌های تشخیص فونت، بازشناسی زبان‌های مختلف و قابلیت حفظ قالب‌بندی سند در نرم‌افزارهای مورد بررسی

ردیف	امکانات و قابلیت‌های	نرم‌افزار صخر	نرم‌افزار گوگل
۱	تشخیص فونت	✓	✓
۲	بازشناسی زبان‌های مختلف	✓	✓
۳	قابلیت حفظ قالب‌بندی نگارش تصویر ورودی (پاراگراف‌بندی، نوع و اندازه فونت‌ها، جدول‌بندی، ستون‌بندی، و ...) در تصویر خروجی	✓	-

نتیجه به دست آمده در جدول ۲، حاکی از آن است که هر دو نرم‌افزار قابلیت بازشناسی فونت‌های گوناگون را دارند؛ چرا که برای سنجش این قابلیت، همان گونه که پیش تر توضیح داده شد، ۱۰ مقاله انتخابی با ۴ فونت مختلف ذخیره و مورد ارزیابی قرار گرفتند. تعداد زبان‌هایی که یک سیستم نویسه‌خوان نوری از آن‌ها پشتیبانی می‌کند، از ویژگی‌های مهم دیگر این سیستم‌هاست. افزون بر آن، بازشناسی زبان انگلیسی با توجه به گستردگی استفاده از این زبان در مقالات و مستندات علمی ضروری است. بنابراین، یک نویسه‌خوان نوری زبان فارسی باید حداقل از زبان انگلیسی نیز پشتیبانی نماید. در بررسی به عمل آمده از این دو نرم‌افزار مشخص شد که علاوه بر زبان فارسی، نرم‌افزار «صخر» از سیزده زبان و نرم‌افزار «گوگل» از بیست و دو زبان رایج دنیا پشتیبانی می‌کنند. در بیشتر آثار تولیدشده علاوه بر متن، نمودار، جدول و تصویر نیز وجود دارد و طرح‌بندی سیستم‌های نرم‌افزاری نویسه‌خوان نوری نیز باید در بازشناسی محتویات این

متون در خروجی کاملاً مشابه باشد. در نرم‌افزارهای «صخر» و «گوگل»، مشابهت طرح‌بندی تصویر (متن، نمودار و جدول) با خروجی تولیدشده مورد بررسی قرار گرفت. در نرم‌افزار «صخر» این قابلیت به‌طور تقریبی وجود دارد، ولی در نرم‌افزار «گوگل» تنها متن، قابل بازشناسی بوده و سایر موارد تشخیص داده نشدند.

۶-۳. قابلیت‌های پردازش گروهی فایل‌ها، دسترسی و کار با دستگاه‌های پوشگر در نرم‌افزارهای مورد بررسی چگونه است؟

نرم‌افزار باید قابلیت پردازش حجم زیادی از تصاویر و تبدیل آن‌ها به نسخه متنی به‌صورت دسته‌ای را داشته باشد. این مهم در مدت‌زمان صرف‌شده برای تبدیل تصویر به متن قابل جست‌وجو تأثیر مستقیم دارد.

جدول ۳. قابلیت‌های پردازش گروهی فایل‌ها، دسترسی و کار با دستگاه‌های پوشگر در نرم‌افزارهای مورد بررسی

ردیف	امکانات و قابلیت‌های	نرم‌افزار صخر	نرم‌افزار گوگل
۱	قابلیت پردازش گروهی فایل‌ها	بله - نامحدود	بله - محدود
۲	قابلیت استفاده	آسان	آسان
۳	قابلیت دسترسی	هزینه‌ای	رایگان
۴	قابلیت کار با دستگاه‌های مختلف پوشگر	✓	مستقل از دستگاه

همان‌طور که در جدول ۳، مشاهده می‌شود، قابلیت‌های پردازش گروهی فایل‌ها در نرم‌افزارهای مورد بررسی در نرم‌افزار «صخر» به‌صورت نامحدود وجود دارد، ولی نرم‌افزار «گوگل» دارای محدودیت بوده و در هر نوبت امکان پردازش حداکثر بیست صفحه را دارد.

از دیگر قابلیت‌های مهم در نرم‌افزارهای «آسی آر»، دسترسی آسان و کار با دستگاه‌های پوشگر است. در این پژوهش این قابلیت‌ها در نرم‌افزارهای «صخر» و «گوگل» مورد بررسی قرار گرفت. قابلیت استفاده در هر دو نرم‌افزار آسان، ولی امکان دسترسی به نرم‌افزار «صخر» همراه با هزینه است، در حالی که دسترسی به نرم‌افزار «گوگل» رایگان است. همچنین، در نرم‌افزار «صخر» قابلیت کار با دستگاه‌های مختلف پوشگر وجود دارد و در نرم‌افزار «گوگل» این قابلیت مستقل از دستگاه است.

۴-۶. توانایی نرم‌افزارهای مورد بررسی در پیش‌پردازش، جداسازی و تشخیص تصاویر و شناسایی طرح‌بندی سند و ارائه طرح‌بندی در خروجی در نرم‌افزارهای مورد بررسی چگونه است؟

بسیاری از اسناد به مرور زمان کیفیت خود را از دست می‌دهند. در نتیجه، قبل از بازشناسی سند نیاز است که پیش‌پردازش‌هایی بر روی سند انجام گرفته و کیفیت آن را بهبود داد. پیش‌پردازش تصویر سند شامل بخش‌های متفاوتی مانند حذف نویز، رفع کج‌شدگی و آستانه‌یابی برای تبدیل آن به یک تصویر است. در جدول ۴، سه توانایی مهم نرم‌افزارهای «اُسی آر» مورد مطالعه، از دید کاربران بررسی شده است.

جدول ۴. توانایی نرم‌افزارهای مورد بررسی در پیش‌پردازش، جداسازی و تشخیص تصاویر و شناسایی طرح‌بندی سند و ارائه طرح‌بندی در خروجی

ردیف	قابلیت‌ها	نرم‌افزار صخر	نرم‌افزار گوگل
۱	توانایی انجام پیش‌پردازش و بهبود کیفیت تصاویر اسناد	با ابزار تنظیمات متنوع	به صورت داخلی و بدون وجود تنظیمات قابل توجه
۲	توانایی در جداسازی و تشخیص تصاویر، اشکال و نمودارها	✓	-
۳	توانایی شناسایی طرح‌بندی سند و ارائه طرح‌بندی در خروجی	✓	-

در بررسی قابلیت‌های نرم‌افزارها یکی از مواردی که حائز اهمیت است، توانایی انجام پیش‌پردازش و بهبود کیفیت تصاویر اسناد است. در ارزیابی دو نرم‌افزار «صخر» و «گوگل»، نرم‌افزار «صخر» با ابزارهایی برای تنظیمات متنوع این قابلیت را داراست، اما در نرم‌افزار «گوگل»، این روند به صورت داخلی و بدون وجود تنظیمات قابل توجه انجام می‌شود. توانایی در جداسازی و تشخیص تصاویر، اشکال و نمودارها در نرم‌افزارهای نویسه‌خوان نوری بسیار مهم است. در برخی اسناد، تصاویر و اشکال در قیاس با متن دارای اهمیت اطلاعاتی بیشتری هستند. در نتیجه، یک نرم‌افزار نویسه‌خوان نوری باید توانایی جداسازی و تشخیص تصاویر، اشکال و نمودارها را داشته باشد. در دو نرم‌افزار مورد بررسی نرم‌افزار «صخر» این توانمندی را داشته و بر خلاف آن، نرم‌افزار «گوگل» این قابلیت را ندارد.

اغلب نویسندگان نوشته خود را به بخش‌هایی تقسیم می‌کنند. این بخش‌ها از لحاظ

نوع اطلاعات متفاوت خواهند بود (متن، نقاشی، تصویر، نمودار، ...). این اجزاء به‌ترتیبی خاص در کنار یکدیگر قرار می‌گیرند که به آن طرح‌بندی ۱ سند گویند. این طرح‌بندی برای آن است که نویسنده مفهوم مورد نظر خود را به شکلی مناسب‌تر به خواننده انتقال دهد. در تبدیل نسخه کاغذی یک سند به نسخه الکترونیکی آن هدف اصلی کاربر ثابت ماندن محتویات است. بنابراین، هر نرم‌افزار بازنشاسی نوری حروف به بخشی برای تحلیل محتویات سند و شکستن آن به نواحی مختلف نیاز داشته و لازم است پس از قطعه‌بندی سند به ناحیه‌های مختلف، در نسخه الکترونیکی سند نیز همان نواحی با ساختار اولیه ایجاد گردد (دبیرخانه شورای عالی اطلاع‌رسانی ۱۳۸۸). در دو نرم‌افزار مورد بررسی، نرم‌افزار «صخر» از این امکان برخوردار است، ولی نرم‌افزار «گوگل» فاقد این امکان است.

۵-۶. امکانات و قابلیت‌های انتخاب، نوع فرمت، تصحیح املائی، بازنشاسی، پیکربندی و تعامل با دستگاه‌های ورودی در نرم‌افزارهای مورد بررسی چگونه است؟
در جدول ۵، هفت مورد از امکانات و قابلیت‌های موجود در نرم‌افزارهای نویسه‌خوان نوری حائز اهمیت آورده شده است.

جدول ۵. امکانات و قابلیت‌های انتخاب، نوع فرمت، تصحیح املائی، بازنشاسی، پیکربندی و تعامل با دستگاه‌های ورودی در نرم‌افزارهای مورد بررسی

ردیف	قابلیت‌ها	نرم‌افزار صخر	نرم‌افزار گوگل
۱	امکان انتخاب بخشی از صفحه توسط کاربر به‌منظور پردازش و ارائه خروجی	✓	-
۲	نوع فرمت‌ها در پذیرش ورودی	همه انواع تصویری رایج	همه انواع تصویری رایج
۳	نوع فرمت‌ها در ارائه خروجی	متن ساده، فایل ورد، اچ‌تی‌ام‌ال فقط متنی	-
۴	امکانات تصحیح املائی خودکار	✓	✓
۵	امکانات یادگیری و شخصی‌سازی	✓	-
۶	امکانات پیکربندی	متنوع	محدود
۷	نحوه تعامل با دستگاه‌های ورودی	به‌صورت مستقیم از طریق رابط‌های ISIS و TWAIN	غیر قابل اتصال

امکان انتخاب بخشی از صفحه توسط کاربر به منظور پردازش و ارائه خروجی در مواردی به کار می‌رود که کاربر فقط به بخشی از یک سند نیاز دارد. بر اساس اطلاعات جدول ۵، در بررسی انجام شده نرم‌افزار «صخر» این قابلیت را دارا بوده و نرم‌افزار «گوگل» فاقد این قابلیت است.

نوع قالب‌های تصویری (مانند BMP، JPEG، TIF) که نرم‌افزار بتواند بپذیرد و آن را پردازش کند نیز اهمیت زیادی دارد. هر دو نرم‌افزار مورد بررسی همه انواع تصاویر رایج را پردازش می‌کنند.

نرم‌افزار نویسه‌خوان باید بتواند پس از پردازش ورودی، خروجی تهیه‌شده را ذخیره کند. از آنجا که نوع فایل ورودی متنوع است، بنابراین، نرم‌افزار نویسه‌خوان باید بتواند نوع فایل خروجی مناسب برای ذخیره کردن خروجی را تشخیص دهد و پس از پردازش ورودی، نتایج تهیه‌شده را به قالب گفته‌شده تبدیل کرده و در محلی که کاربر مشخص خواهد کرد، ذخیره کند. تنوع قالب‌های خروجی پشتیبانی‌شده از سوی نویسه‌خوان‌های نوری بسیار مهم است، زیرا کاربران اغلب پس از تهیه خروجی، در صورتی که خروجی در قالب مورد نیاز نباشد، مجبور به تبدیل خروجی به قالب مناسب خواهند شد. قالب‌های متداول برای ذخیره فایل‌های متنی عبارت‌اند از TXT، RTF، XML و HTML (دیرخانه شورای عالی اطلاع‌رسانی ۱۳۸۸). همان‌طور که در جدول ۵، آورده شده، نرم‌افزار «صخر» این قابلیت را دارد که خروجی‌های خود را به صورت متن ساده (TXT)، فایل «ورد» و «اچ‌تی‌ام‌ال» ارائه دهد، ولی نرم‌افزار گوگل فقط به صورت TXT خروجی می‌دهد.

در نرم‌افزارهای نویسه‌خوان مؤلفه‌هایی برای تصحیح املائی وجود دارد. به این دلیل که خروجی هر نرم‌افزار بسته به دقت آن دارای تعدادی غلط املائی است، در نرم‌افزارهای «صخر» و «گوگل» امکانات تصحیح املائی خودکار وجود دارد. در فرایند بازشناسی، اغلب کاربران به امکانات یادگیری و شخصی‌سازی نیاز دارند؛ چرا که می‌خواهند حجم بزرگی از اطلاعات با ساختار و ویژگی‌های یکسان را به نسخه الکترونیکی تبدیل کنند (مثلاً یک کتاب، یک گزارش و ...). طبیعی است که چنین متن‌هایی ویژگی‌های خاصی دارند (مانند قلم خاص، طرح‌بندی یکسان و ...). این ویژگی‌ها اطلاعات بازرشی در خود خواهند داشت که می‌توان از آن در بازشناسی و شخصی‌سازی متن بهره برده و دقت سیستم را افزایش داد. اغلب سیستم‌های موفق بازشناسی نوری حروف امکاناتی دارند که کاربر به کمک آن می‌تواند اطلاعات مفیدی در مورد نوع متن ورودی به سیستم ارائه

دهد و سیستم از این اطلاعات برای بازشناسی بهتر متن بهره‌مند شود (دبیرخانه شورای عالی اطلاع‌رسانی ۱۳۸۸). در بررسی انجام‌شده، نرم‌افزار «صخر» از این امکان برخوردار بوده، ولی نرم‌افزار «گوگل» فاقد این امکان است.

نرم‌افزار برای اجرای بهینه و درست الگوریتم‌های مقادیر آستانه پارامترهایی دارد که می‌بایست مقدار آن‌ها به صورت پیش فرض تعیین شده یا توسط کاربر تنظیم شود. مسائل دیگری نیز وجود دارد؛ مانند نحوه استفاده از منابع نرم‌افزار. برای مثال، نحوه مدیریت حافظه (مانند ذخیره‌سازی ویژگی‌ها، قالب‌ها، قلم‌ها و ... در حافظه یا دیسک جانبی) یا نحوه انجام پردازش‌ها یا استفاده از پردازش که مربوط به پیکربندی بوده و نرم‌افزار نویسه‌خوان می‌بایست برای مقداردهی آن‌ها بازه مناسبی تعیین کرده و مقادیر پیش فرض مناسب برای کاربر نا آشنا با جزئیات نرم‌افزار را در نظر گیرد (دبیرخانه شورای عالی اطلاع‌رسانی ۱۳۸۸). در نرم‌افزار «صخر» این امکانات به صورت متنوع وجود دارد، ولی در گوگل محدود است.

آخرین موردی که برای ارزیابی این دو نرم‌افزار بازشناسی نوری مد نظر قرار گرفته، نحوه تعامل با دستگاه‌های ورودی است. یک سیستم بازشناسی مناسب باید توانایی کار با انواع مختلفی از اسکنرها و مدل‌های مختلف آن را داشته باشد و دقت آن با توجه به نوع دستگاه تغییر چندانی نکند. در سیستم‌های جدید می‌توان از دوربین‌های دیجیتال نیز به عنوان دستگاه ورودی استفاده کرد. در نرم‌افزار «صخر» این امکان با رابط‌های TWAIN و ISIS فراهم شده است، ولی در نرم‌افزار «گوگل» غیر قابل اتصال است.

۷. بحث و نتیجه‌گیری

از آنجا که در متن‌های لاتین، کلمات از حروف مجزا تشکیل شده‌اند و فاصله طبیعی در بین کلمات وجود دارد، تشخیص محدوده کلمه آسان بوده، ولی در متن‌های فارسی و عربی که اغلب کلمات متصل هستند، تفکیک و تشخیص کلمات بسیار دشوار است. به همین دلیل، با این که سال‌ها از عمر نرم‌افزارهای نویسه‌خوان نوری در دنیا می‌گذرد، با توجه به مشکلاتی (از جمله وجود نقطه، حروف چسبیده و غیر چسبیده، و تفاوت‌های نوشتاری مانند سر هم و جدانویسی برخی از واژه‌ها که در نگارش خط فارسی وجود دارد)، هنوز نرم‌افزار نویسه‌خوان نوری فارسی که بتواند تمام مشکلات موجود را حل کند، به بازار نیامده است. یکی از نرم‌افزارهای معتبر در این حوزه نرم‌افزار عربی «صخر»

است که می‌تواند از زبان فارسی نیز پشتیبانی کند. شرکت «گوگل» نیز در سال‌های اخیر اقدام به توسعه نرم‌افزار نویسه‌خوان نوری فارسی در «گوگل‌درایو» نموده است. در این مقاله به ارزیابی این دو نرم‌افزار نویسه‌خوان نوری پرداخته شد.

از جمله مهم‌ترین نتایج به‌دست آمده این که می‌توان گفت نرم‌افزار «صخر» از نظر دقت بازشناسی و صحت تشخیص حرف و کلمه در جایگاه پایین‌تری از نرم‌افزار «گوگل» قرار دارد. همچنین، میزان خطا در نرم‌افزار «صخر» نسبت به نرم‌افزار «گوگل» بالاتر است. «کانونگو، مارتون و بولبول» در پژوهش خود در مقایسه درصد خطای نرم‌افزار «صخر» با «اومنی‌پیج» اذعان داشتند که «صخر» نسبت به «اومنی‌پیج» ۱۵/۴۷ درصد کمتر دچار خطا می‌شود (Kanungo, Marton and Bulbul 1998). این میزان در پژوهش حاضر کمتر شده است. سرعت «اُسی‌آر» در نرم‌افزار «صخر» در هر ۲ ثانیه ۱ صفحه است، ولی در نرم‌افزار «گوگل» به دلیل آنلاین بودن سیستم، میزان سرعت قابل اندازه‌گیری نبود و دلیل اصلی آن آنلاین بودن نرم‌افزار «گوگل» و آفلاین بودن نرم‌افزار «صخر» است.

هر دو نرم‌افزار قابلیت بازشناسی فونت‌های گوناگون را دارند. نرم‌افزار «صخر» علاوه بر زبان فارسی، از سیزده زبان، و نرم‌افزار «گوگل» از بیست‌ودو زبان رایج دنیا پشتیبانی می‌کند. «ساجدی» در پژوهش خود به پیشرفت زیاد نرم‌افزارهای نویسه‌خوان اشاره می‌کند (Sajedi 2016). در نرم‌افزارهای «صخر»، قابلیت مشابهت طرح‌بندی تصویر (متن، نمودار و جدول) با خروجی تولیدشده مورد بررسی قرار گرفت. در نرم‌افزار «صخر» این قابلیت، به‌طور تقریبی وجود دارد، ولی در نرم‌افزار «گوگل» تنها متن قابل بازشناسی بوده و سایر موارد تشخیص داده نشدند. «دیمان و سینق» در پژوهش خود به این نکته به‌عنوان ضعف نویسه‌خوان‌ها اشاره داشتند (Dhiman and Singh 2013).

در نرم‌افزار «صخر» قابلیت‌های پردازش گروهی فایل‌ها به‌صورت نامحدود وجود دارد، ولی نرم‌افزار «گوگل» دارای محدودیت بوده و در هر نوبت امکان پردازش حداکثر بیست صفحه را دارد. در این مورد نتایج پژوهش حاضر با نتایج پژوهش (Kanungo, Marton and Bulbul 1999) در یک راستا قرار دارند. قابلیت استفاده در هر دو نرم‌افزار آسان است. این در حالی است که امکان دسترسی به نرم‌افزار «صخر» همراه با هزینه بوده، ولی دسترسی به نرم‌افزار «گوگل» رایگان است. همچنین، در نرم‌افزار «صخر» قابلیت کار با دستگاه‌های مختلف پوشگر وجود دارد و در نرم‌افزار «گوگل» این قابلیت مستقل از دستگاه موجود است. امکان انجام پیش‌پردازش و بهبود کیفیت تصاویر اسناد در نرم‌افزار

«صخر» با ابزارهایی برای تنظیمات متنوع فراهم بوده، اما در نرم‌افزار «گوگل» این روند به‌صورت داخلی و بدون وجود تنظیمات قابل توجه انجام می‌شود. نتایج این بخش از پژوهش نیز با نتایج ارائه‌شده در پژوهش (Kanungo, Marton and Bulbul 1998) همخوانی دارد. توانایی در جداسازی و تشخیص تصاویر و توانایی شناسایی طرح‌بندی سند و ارائه طرح‌بندی در خروجی در نرم‌افزار «صخر» وجود دارد، در حالی که نرم‌افزار «گوگل» فاقد این توانمندی‌هاست.

هر دو نرم‌افزار مورد بررسی همه‌انواع تصاویر رایج را پردازش می‌کنند. نرم‌افزار «صخر» این قابلیت را دارد که خروجی‌های خود را به‌صورت متنی، «ورد» و «اچ‌تی‌ام‌ال» ارائه دهد، ولی نرم‌افزار «گوگل» فقط به‌صورت متنی خروجی می‌دهد. در هر دو نرم‌افزار امکانات تصحیح املایی خودکار وجود دارد. قابلیت یادگیری و شخصی‌سازی امکانات، پیکربندی، و نحوه تعامل با دستگاه‌های ورودی در نرم‌افزار «صخر» وجود دارد، ولی نرم‌افزار «گوگل» فاقد این قابلیت است.

در مجموع، می‌توان این‌گونه بیان داشت که هر یک از این دو نرم‌افزار قابلیت‌های مهم و خوبی دارند که برای استفاده کتابخانه‌ها و مراکز اطلاع‌رسانی بسیار ضروری است. از سوی دیگر، نقاط ضعف هر یک نیز مانع استفاده بهینه از آن‌هاست.

۸. پیشنهادهای کاربردی و پژوهشی

- ◇ از آنجا که نرم‌افزار «صخر» امکان رفع مشکلات برای هر مرکز و سفارشی‌سازی «اُسی‌آر» را دارد، پیشنهاد می‌شود کتابخانه‌ها و مراکز اطلاع‌رسانی که توانایی پرداخت هزینه برای خرید را دارند، از این نرم‌افزار استفاده کنند؛
- ◇ پیشنهاد می‌شود کتابخانه‌ها و مراکز اطلاع‌رسانی که توانایی پرداخت هزینه برای خرید نرم‌افزار «صخر» را ندارند، برای ایجاد امکان جست‌وجوی درون‌متنی در منابع دیجیتال از نرم‌افزار «گوگل» استفاده کنند، اگرچه استفاده از «گوگل» امکان رفع مشکلات هر مرکز و سفارشی‌سازی «اُسی‌آر» را ندارد.

۹. پیشنهاد برای پژوهش‌های آتی

- ◇ با توجه به سرعت پیشرفت در حوزه‌های فناوری اطلاعات و ارائه امکانات بیشتر در نسخه‌های جدید در نرم‌افزارهای گوناگون برای جلب رضایت کاربران، پیشنهاد

می‌شود امکانات و قابلیت‌های این دو نرم‌افزار نویسه‌خوان نوری با یکدیگر یا با نرم‌افزارهای نویسه‌خوان نوری جدید هر سال مورد ارزیابی قرار گیرند.

فهرست منابع

- آگاهی، حامد، و مرضیه صالحی. ۱۳۹۵. مروری بر سیستم‌های بازشناسی ارقام دست‌نویس فارسی. اولین کنفرانس بین‌المللی دستاوردهای نوین پژوهشی در مهندسی برق و کامپیوتر. تهران.
- ایرانپور مبارکه، مجید، و علیرضا احمدی‌فرد. ۱۳۹۵. ارائه خصیصه‌های خاص زبان فارسی جهت بازیابی و بازشناسی کلمات تصویری فارسی با استفاده از تعبیه برچسب. مجله ماشین‌بینایی و پردازش تصویر ۳ (۱): ۴۵-۵۴.
- حریری، نجلا. ۱۳۸۵. اصول و روش‌های پژوهش کیفی. تهران: دانشگاه آزاد اسلامی، واحد علوم و تحقیقات.
- خسروی، حسین، و احسان‌اله کیپر. ۱۳۸۸. ارزیابی روش‌های بازشناسی متون فارسی بر مبنای شکل کلی زیر کلمات. نشریه مهندسی برق و مهندسی کامپیوتر ایران ۷ (۴): ۲۶۷-۲۸۰.
- دبیرخانه شورای عالی اطلاع‌رسانی. ۱۳۸۸. پژوهشنامه نویسه‌خوان نوری فارسی (ocr). تهران: دبیرخانه شورای عالی اطلاع‌رسانی.
- زند، محسن. ۱۳۹۱. تشخیص حروف چاپی فارسی با استفاده از روش ترکیبی. همایش منطقه‌ای علوم کامپیوتر، مهندسی کامپیوتر و فناوری اطلاعات. دو رود.
- سلیمانی، سید محمد، و هادی حسینی. ۱۳۹۴. یک رویکرد مبتنی بر مدل برای شناسایی و اصلاح نویسه‌های فارسی و عربی مستقل از متن آفلاین. دومین کنگره سراسری فناوری‌های نوین ایران با هدف دستیابی به توسعه پایدار. تهران.
- فرامرزی، اسماعیل. ۱۳۸۴. بازشناسی نوری حروف: مروری بر مباحث نظری و ملاحظات کاربردی با تأکید بر مسائل خاص زبان فارسی. پژوهشنامه پردازش و مدیریت اطلاعات ۲۰ (۳-۴): ۳۳-۶۱.
- هنروران، فرامرز. ۱۳۷۵. اسی آر و کاربردهای آن در کتابخانه‌ها و مراکز اطلاع‌رسانی. تحقیقات اطلاع‌رسانی و کتابخانه‌های عمومی ۲۰ (۶): ۵۰-۵۳.

References

- Dhiman, Shivani, and A Singh. 2013. Tesseract vs gocr a comparative study. *International Journal of Recent Technology and Engineering* 2 (4): 80.
- Kanungo, Tapas, Gregory A. Marton, and Osama Bulbul. 1998. Performance evaluation of two Arabic OCR products. Proceedings of SPIE. Paper presented at the International Society for Optical Engineering.
- _____. 1999. OmniPage vs. Sakhr: Paired model evaluation of two Arabic OCR products. Proceedings of SPIE. Paper presented at the International Society for Optical Engineering.
- Sajedi, Hedieh. 2016. Handwriting recognition of digits, signs, and numerical strings in Persian. *Computers & Electrical Engineering*. 52-65 :49 .

Shayegan, Mohammad Amin, and Chee Seng Chan. 2012. A new approach to feature selection in handwritten farsi/ arabic character recognition. *Advanced Computer Science Applications and Technologies (ACSAT). International Conference on Advanced Computer Science Applications and Technologies*. Kuala Lumpur, Malaysia.

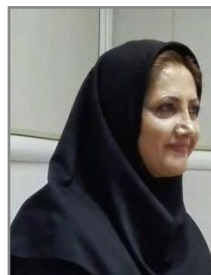
مهرناز خراسانچی

دارای مدرک دکتری در رشته علم اطلاعات و دانش‌شناسی است. ایشان هم‌اکنون پژوهشگر حوزه تفکر پلتفرمی در کتابخانه‌ها و مراکز اطلاع‌رسانی است. تفکر پلتفرمی، کتابداری کودک، کتابخانه‌های دیجیتال، کتابخانه‌های عمومی، از جمله علایق پژوهشی وی است.



طیبه شه میرزادی

دارای مدرک دکتری در رشته علم اطلاعات و دانش‌شناسی است. ایشان هم‌اکنون استادیار مرکز فناوری اطلاعات و اطلاع‌رسانی کشاورزی در سازمان تحقیقات، آموزش و ترویج کشاورزی است. علم‌سنجی، مدیریت اطلاعات، فراهم‌آوری اطلاعات، و کتابخانه‌های دیجیتال از جمله علایق پژوهشی وی است.



پژوهشگاه علوم انسانی و مطالعات فرهنگی
رتال جامع علوم انسانی



شروېشگاه علوم انساني و مطالعات فرهنگي
پرتال جامع علوم انساني