

به نام خدا



وزارت علوم، تحقیقات و فناوری

پژوهشگاه علوم و فناوری اطلاعات ایران

(ایرانداک)

خلاصه گزارش طرح پژوهشی سازمانی

عنوان طرح پژوهشی

دسته‌بندی تصویرها در مدارک پایگاه اطلاعات علمی ایران (گنج) بر اساس یادگیری عمیق

تاریخ شروع طرح پژوهشی	۱۴۰۰/۱۰/۲۲	تاریخ پایان طرح پژوهشی	۱۴۰۲/۰۱/۲۰
نام و نام خانوادگی مجری طرح پژوهشی	آزاده فخرزاده		
نام و نام خانوادگی همکاران طرح پژوهشی	امیر حسین صدیقی		

درباره طرح پژوهشی

تصویرها در اسناد علمی حاوی مطالب مهمی هستند. نویسندگان معمولاً خلاصه‌ای از روش پیشنهادشده و نتایج مقاله را در قالب نمودار نمایش می‌دهند. تصویرها اطلاعات زیادی را در زمان کوتاهی منتقل می‌کنند. چنانچه موتورهای جستجو، تصویرهای مربوط به جستجوی کاربر را در کنار فراداده‌هایی متنی در اختیار او قرار دهند، کاربر با سرعت بیشتری می‌تواند به اطلاعات موردنظر دسترسی پیدا کند. به همین دلیل امروزه تصویرها به‌عنوان یکی از منابع اطلاعاتی مهم در نظر گرفته می‌شوند و بحث بازیابی آن‌ها از پایگاه‌های داده از موضوعات مهم و به‌روز است. در «گنج» امکان جستجو بر اساس یک عبارت متنی و بازیابی و نمایش نتایج جستجو در قالب فراداده‌های متنی (عنوان، چکیده، کلیدواژه، پدیدآور، سال انتشار) وجود دارد. لیکن تصویرهای موجود در اسناد «گنج» بازیابی نمی‌شوند. فراهم کردن چنین امکانی در «گنج» به‌عنوان یک ارزش افزوده می‌تواند آن را از موتورهای جستجوی مشابه متمایز سازد. بازیابی اطلاعات از تصویرها به دو روش بافت محور و محتوا - محور امکان‌پذیر است. در روش محتوا - محور، محتوای بصری تصویر برای بازیابی در نظر گرفته می‌شود. به دلیل پیچیدگی‌های تصویرهای علمی، جهت استخراج اطلاعات از محتوای تصویرها و استفاده از روش‌های محتوا - محور، ابتدا باید آن‌ها را دسته‌بندی کرد. در این تحقیق، ابتدا ۹۱۴۴ تصویر به‌صورت تصادفی در شش حوزه مختلف از رساله‌ها و پایان‌نامه‌های گنج انتخاب شد و توسط خبرگان برچسب زده شد. تصاویر به یازده دسته شامل نمودار میله‌ای، تصویر طبیعی، نمودار دایره‌ای، نمودار ساختارمند، رسم فنی، منحنی x-y، نقشه‌ها تصاویر اسلاید، جدول‌ها، نمودارهای جعبه‌ای، نمودارهای سه‌بعدی دسته‌بندی شدند. با توجه به کوچک بودن حجم داده آموزشی، روش یادگیری عمیق بر اساس شبکه موبایل نت به کار گرفته شد. از آنجاییکه داده آموزشی به شدت نامتقارن بود، از روش اعتبارسنجی متقابل طبقه‌ای و تابع زیان آنتروپی متقاطع در آموزش شبکه استفاده شد. مدل معرفی شده با دقت ۰٫۹۰ قابلیت دسته‌بندی تصاویر موجود در گنج را دارد.

روش پژوهش

روش پژوهش برای بررسی انواع روش‌ها و الگوریتم‌های موجود، روش مطالعات کتابخانه‌ای است. با مطالعه تحقیقات پیشین و بررسی چالش‌های مسئله پیش رو، روش جدیدی پیشنهاد می‌شود. روش‌های اخیر پیشنهاد شده برای دسته‌بندی تصویرهای علمی بر اساس یادگیری عمیق است. براین اساس در این تحقیق نیز الگوریتم‌های یادگیری عمیق بررسی می‌شود و الگوریتم مناسب پیشنهاد می‌شود. سپس روش پیشنهادی بر روی مجموعه‌ای از داده‌ها آزمایش می‌شود تا صحت و دقت، و خطای احتمالی آن بررسی گردد و در صورت نیاز بهبودهایی روی آن صورت گیرد؛ بنابراین، آماده‌سازی این مجموعه داده برای انجام آزمایش و اندازه‌گیری دقت روش هم قسمتی از فعالیت‌های این پژوهش محسوب می‌شود

یافته‌های طرح پژوهشی

در این طرح یک مجموعه داده از تصاویر گنج تشکیل شد و الگوریتم یادگیری عمیق برای دسته‌بندی تصاویر آموزش داده شد. این الگوریتم برای پیش‌بینی دسته تصاویر موجود در اسناد گنج می‌تواند استفاده شود. یک مجموعه داده شامل حدود ۹۱۴۴ نمونه از تصاویر موجود در گنج توسط خبرگان برچسب‌گذاری شدند، یازده دسته در نظر گرفته شد. این یازده دسته شامل نمودار میله‌ای، تصویر طبیعی، نمودار دایره‌ای، نمودار ساختارمند، رسم‌فنی، منحنی $X-Y$ ، نقشه‌ها، تصاویر اسلاید، جدول‌ها، نمودارهای جعبه‌ای و نمودارهای سه‌بعدی است. سپس با استفاده از شبکه موبایل نت نسخه دو (MobileNetV2) مدل دسته‌بندی تصاویر ساخته و آموزش داده شد. موبایل نت به دلیل اینکه تعداد پارامترهای کمتری نسبت به سایر شبکه‌های یادگیری عمیق دارد در مسائل با حجم داده آموزشی کم، مانند داده آموزشی این طرح، کاربرد دارد. دقت مدل پیشنهادی روی داده آموزشی ۰,۹ است. مدل آموخته شده می‌تواند برای پیش‌بینی دسته داده‌های تصویری گنج به کار رود.

نتیجه‌گیری و پیشنهادها

در این طرح یک مدل با دقت بالا برای دسته‌بندی تصاویر موجود در گنج ارائه شد. تصاویر آزمایشی حدود نه هزار تصویر بودند و اعمال مدل روی داده آزمایشی بزرگتر اعتبار آن را بیشتر می‌کند. در مرحله بعد می‌توان بر اساس ویژگی‌های هر دسته از تصاویر، از ابزارهای مناسب پردازش تصویر برای استخراج اطلاعات از آنها استفاده کرد و

برچسب‌های توصیف‌کننده به هر تصویر اختصاص داد. همچنین می‌توان یک مدل ارائه داد که ارتباط بین اطلاعات متنی موجود در فایل تصویر، و اطلاعات موجود در تصویر را تشخیص دهد و برچسب‌های دقیق‌تری برای هر تصویر استخراج کند.