

دومین کنفرانس ملی مهندسی برق ایران - دانشگاه آزاد اسلامی واحد نجف آباد - ۵ و ۶ اسفند ماه ۱۳۸۸

استخراج اطلاعات معنایی از تصاویر متنی مبتنی بر طبقه‌بندی

حسین پورقاسم

گروه مخابرات دانشکده برق - دانشگاه آزاد اسلامی واحد نجف آباد

h_pourghasem@iaun.ac.ir

خلاصه: شناسایی و استخراج اطلاعات از تصاویر متنی در اینترنت یکی از مهمترین مسائل مهم سیستم‌های جمع‌آوری اطلاعاتی در حوزه تکنولوژی اطلاعات می‌باشد. در این مقاله، جهت تجزیه و تحلیل تصاویر متنی و دسترسی به محتوای آنها، از یک روش طبقه‌بندی مبتنی بر ناحیه‌بندی تصویر استفاده شده است. در این الگوریتم به کمک یک روش ناحیه‌بندی دو مرحله‌ای، نواحی تصویر شناسایی شده، سپس به کمک یک ساختار سلسله مراتبی طبقه‌بندی، نوع ناحیه از لحاظ متنی یا عکسی (غیر متنی) بودن مشخص می‌گردد. در ادامه با تعریف با ارزش بودن متن یک تصویر متنی، سعی می‌شود تصویر متنی در یکی از دو گروه معنایی با ارزش و بی‌ارزش دسته‌بندی گردد. الگوریتم پیشنهادی بر روی یک پایگاه داده تصاویر متنی و غیرمتنی که از تصاویر موجود در اینترنت فراهم شده است، مورد ارزیابی قرار می‌گیرد. نتایج آزمایشها بیانگر کارایی روش پیشنهادی در طبقه‌بندی معنایی تصاویر براساس تعریف کاربر از با ارزش و بی‌ارزش بودن تصاویر متنی است. الگوریتم پیشنهادی، صحت طبقه‌بندی ۹۸/۸٪ برای طبقه‌بندی تصاویر متنی با ارزش از بی‌ارزش را فراهم کرده است.

کلمات کلیدی: استخراج اطلاعات، طبقه‌بندی، تصاویر متنی و غیر متنی، ارزشمندی اطلاعاتی.

۱ - مقدمه

رشد فزاینده اسناد کاغذی دیجیتالی در مجاری اینترنتی و سامانه‌های خودکارسازی اداری، نیاز به بایگانی و دسترسی آنها جهت استخراج و داده‌کاوی اطلاعات را به یک نیاز مهم تبدیل کرده است. با توجه به زمان بر بودن این فرآیند بصورت دستی (توسط کاربر)، ارائه ابزار خودکاری که بتواند این نیاز را مرتفع کند، ضروری به نظر می‌رسد. بنابراین در یک سامانه خودکار جستجو و بازیابی اطلاعات متنی، تجزیه و تحلیل اسناد متنی جهت دسترسی به اطلاعات این اسناد امری اجتناب‌ناپذیر تلقی می‌شود [۱]. تجزیه و تحلیل اسناد اسکن شده، مقدمه‌ای بر دسترسی به اطلاعات معنایی و مفهومی نهفته در تصویر متنی است. این اطلاعات سطح بالا جهت طبقه‌بندی معنایی تصاویر می‌تواند ابزار نیرومندی در اختیار طراحان سیستم‌های جمع‌آوری اطلاعات قرار دهد. اسناد به سه روش عمده تجزیه و تحلیل می‌شوند [۲]:

۱- تجزیه و تحلیل بر اساس محتویات سند: در این دسته روش صرفنظر از اینکه چطور اسناد ایجاد شده‌اند، براساس محتویاتشان به دو دسته تقسیم می‌شوند: طبیعی (مانند عکس صورت و تصاویر

ماهواره‌ای) و نمادین (مانند کارهای چاپی، فرم‌های اداری و طراحی‌های مهندسی) [۲].

۲- تجزیه و تحلیل بر اساس لایه‌بندی سند: در این دسته روش، لایه‌بندی سند مورد بررسی قرار می‌گیرد. این دسته روش نیز به نوبه خود به دو دسته تقسیم می‌شود: تجزیه و تحلیل هندسی (فیزیکی) و تجزیه و تحلیل منطقی. تجزیه و تحلیل هندسی، خصوصیات هندسی است که مناطق همگون مانند متن، جدول، تصویر، نقاشی و ... را دسته‌بندی می‌کند. در مقابل، تجزیه و تحلیل منطقی، به نوع صفحه‌بندی می‌پردازد و برچسب‌های موجود در صفحه مانند عنوان، علامت تجاری، پاورقی، حروف بزرگ و کوچک و ... را مشخص می‌کند [۳، ۱].

۳- تجزیه و تحلیل بر اساس متن: در این تجزیه و تحلیل، سند به دو قسمت متنی و غیرمتنی (گرافیکی) تقسیم می‌شود. برای دسته‌بندی و جداسازی نواحی متنی از نواحی گرافیکی، از فیلترهای موجود در پردازش تصویر استفاده می‌شود [۴-۵].

اکثر کارهایی که اخیراً انجام شده است، تجزیه و تحلیل سند را براساس ناحیه‌بندی تصاویر سندی به نواحی متنی و گرافیکی موجود

در تصویر انجام داده اند (دسته روش سوم). در این تحقیقات ناحیه بندی معمولاً به یکی از سه روش بالا به پایین، پایین به بالا و مبتنی بر بافت انجام شده است [۶،۷].

روش هایی که به صورت پایین به بالا عمل می کنند، از اجزای کوچک تصویر (پیکسل) استفاده کرده و با اتصال این اجزا به هم و تشکیل اجزای بزرگتر برای شناسایی قطعات اصلی موجود در تصویر استفاده می کنند. به عبارت دیگر این روشها، ناحیه بندی را از پیکسل شروع کرده و به ناحیه یا شئی می رسند. اشکال این روش در حساس بودن به اندازه حروف، وضوح سند اسکن شده و فاصله میان حروف است [۴]. از جمله این روش ها می توان به الگوریتم Docstrum اشاره کرد که بر اساس K-نزدیکترین همسایگی از اجزا صفحه عمل می کند [۸]. در روش هایی که به صورت بالا به پایین عمل می کنند، عملیات ناحیه بندی از بزرگترین جزء، یعنی تمام صفحه شروع شده تا به کوچکترین قطعه در صفحه می رسد. اشکال این روش در تشخیص زاویه و انحنای متن است. زیرا محدود به بلوکهای مستطیلی است [۴]. از جمله این روش ها می توان به الگوریتم بازگشتی X-Y Cut [۹] اشاره کرد. در هر مرحله مقادیر واریانس در جهت های افقی و عمودی محاسبه شده و با مقدار آستانه ای که از پیش تعریف شده است، مقایسه می گردد و در انتها نیز ناحیه بندی انجام می گیرد. روش هایی که بافت را برای ناحیه بندی بکار می برند، از ویژگیهای بافت تصویر برای ناحیه بندی استفاده می کنند [۱۰]. در دو روش اول (پایین به بالا، بالا به پایین) نیاز به پیش اطلاعاتی درباره فاصله های افقی و عمودی، مجزا بودن متن ها، عکس ها، گرافیک ها و حتی فرض هایی درباره متن، سبک و اندازه می باشد. با توجه به اشکالات و موارد مطرح شده در مورد روشهای ناحیه بندی، در این مقاله از روش ناحیه بندی مبتنی بر بافت استفاده می شود.

در بسیاری از روشهای ارائه شده قبلی در زمینه طبقه بندی، برای آنکه ویژگیها با اطلاعات محلی در اختیار قرار گیرند تا عملکرد طبقه بندی بهبود یابد، ویژگیها بصورت محلی استخراج می شوند [۱۱]. برای این منظور، تصویر به چندین ناحیه ثابت مثلاً بلوکهای مستطیلی هم اندازه تقسیم شده و ویژگیها از هر بلوک استخراج می شوند. در این مقاله، با استفاده از یک ساختار طبقه بندی سلسله مراتبی مبتنی بر ناحیه بندی تصاویر سندی ارائه شده در [۱]، نواحی متنی و گرافیکی تصویر ورودی مشخص می شود. در این نقشه، از یک الگوریتم ناحیه بندی دو مرحله ای مبتنی بر تبدیل موجک و آستانه گذاری استفاده می شود. در اولین مرحله از الگوریتم ناحیه بندی، از تبدیل موجک [۴-۵،۷،۱۲] برای ناحیه بندی استفاده می شود. سپس با استفاده از یک طبقه بند شبکه عصبی پرسپترون چندلایه (MLP)، برچسب نواحی از نظر متنی بودن یا گرافیکی (غیرمتنی) بودن مشخص می گردد و نواحی متنی شناسایی شده از تصویر کنار گذاشته می شوند. در مرحله دوم با اعمال یک ناحیه بندی ثانویه مبتنی بر آستانه گذاری بر روی باقی مانده تصویر حاصل از ناحیه بندی اولیه، تصویر بصورت کامل ناحیه بندی می شود. در

انتها نیز برچسب نواحی ناحیه بندی شده توسط ناحیه بند دوم توسط دو طبقه بند (MLP و ماشین بردار پشتیبان (SVM)) مشخص می گردد. در هر سه طبقه بند، از ویژگیهای بافت جهت طبقه بندی نواحی استفاده می شود. ویژگیهایی همچون Tamura (شامل زبری، جهت دار بودن و کنتراست)، همبستگی، انرژی، همگنی و آنتروپی مستخرج از ماتریس همخوانی و همچنین دو ویژگی از زیرباندهای تبدیل موجک استخراج می گردد. در ادامه با تعریف الگوی ارزشمندی متن و استخراج ویژگی از این الگو، یک طبقه بند شبکه عصبی آموزش داده می شود تا در مرحله آزمایش تصاویر متنی با ارزش از بقیه تصاویر مجزا گردد.

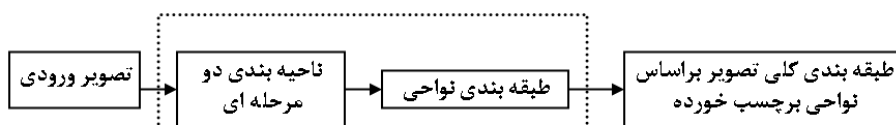
در ادامه این مقاله، در بخش ۲ چارچوب پیشنهادی طبقه بندی تصاویر معرفی می گردد. در این بخش، جزئیات ناحیه بندی دو مرحله ای، استخراج ویژگی، تعریف الگوی ارزشمندی متن و طبقه بندی معنایی تصویر تشریح می شود. در بخش ۳ آزمایشهای انجام شده جهت ارزیابی الگوریتم پیشنهادی آورده شده است. بخش ۴ مروری بر عملکرد الگوریتم پیشنهادی ارائه می دهد. در انتها نیز در بخش ۵ نتیجه گیری از کار انجام شده در این مقاله ارائه می گردد.

۲ - چارچوب پیشنهادی طبقه بندی تصاویر

شکل (۱) روندنمای کلی چارچوب پیشنهادی برای طبقه بندی معنایی تصاویر متنی مبتنی بر طبقه بندی نواحی تصویر را نشان می دهد. در چارچوب پیشنهادی تصویر ورودی، طی دو مرحله ناحیه بندی، به یک سری نواحی با برچسبهای متنی و گرافیکی ناحیه بندی می شود. در مرحله آخر نیز، با آموزش الگوهای تصاویر متنی با ارزش و بی ارزش تصاویر مطلوب و نامطلوب از یکدیگر جدا می شوند. در این مقاله جهت طبقه بندی نواحی مستخرج از تصویر از الگوریتم معرفی شده در [۱] استفاده شده است. روندنمای طبقه بند سلسله مراتبی مبتنی بر ناحیه بندی با جزئیات بیشتر در شکل (۲) نشان داده شده است.

در [۱]، از یک الگوریتم ناحیه بندی دو مرحله ای برای ناحیه بندی تصویر و به دنبال آن طبقه بندی نواحی استفاده شده است. در مرحله اول این الگوریتم، تصویر ورودی با یک روش ناحیه بندی مبتنی بر تبدیل موجک ناحیه بندی می شود. سپس این نواحی با طبقه بند اول، طبقه بندی می شود. نواحی متنی شناسایی شده از تصویر ورودی حذف شده و نواحی گرافیکی در تصویر باقی می ماند تا دوباره و با دقت بیشتر توسط ناحیه بند دوم (ناحیه بند مبتنی بر آستانه گذاری) ناحیه بندی شده و به کمک دو طبقه بند (طبقه بند دوم و سوم)، متنی یا گرافیکی بودن نواحی مشخص می گردد. برای استفاده بهتر از الگوریتم طبقه بندی مبتنی بر ناحیه بندی ارائه شده در [۱]، مرحله استخراج ویژگی از نواحی را تغییر دادیم. بدین صورت که با اضافه کردن یک ویژگی بافت دیگر (ویژگی Tamura) به بردار ویژگی سعی شده است که بردار ویژگی تقویت شود. در ادامه تنها به تشریح ویژگی جدید Tamura پرداخته می شود.

طبقه‌بندی تصویر مبتنی بر ناحیه‌بندی



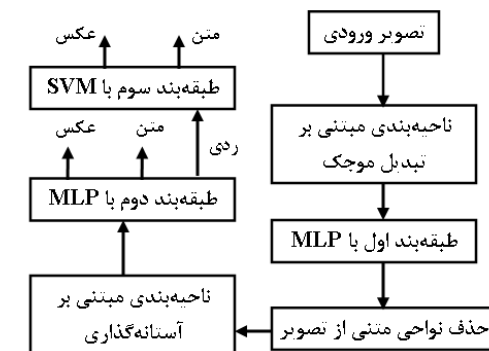
شکل (۱): روندنمای کلی چارچوب پیشنهادی طبقه‌بندی معنایی تصاویر.

ب- یک ناحیه متنی در زیر لوگو که بیانگر عنوان نامه یا طرفین مکاتبه باشد.

ج- یک ناحیه متنی در زیر ناحیه عنوان که بیانگر بدنه نامه می‌باشد.

د- یک ناحیه گرافیکی در زیر ناحیه بدنه که بیانگر مهر یا امضاء نامه می‌باشد.

در موارد ذکر شده فوق دو مسأله بصورت نسبی مطرح می‌گردد. یکی، موقعیت نواحی در صفحه و همچنین موقعیت نسبی نواحی نسبت به یکدیگر است. دیگری، مساحت نسبی هر ناحیه متنی یا گرافیکی می‌باشد. شکل (۳) الگوی نوعی یک متن با ارزش (نامه رسمی- اداری) را نشان می‌دهد.



شکل (۲): روندنمای طبقه‌بند سلسله مراتبی مبتنی بر ناحیه‌بندی [۱].

۱-۲- استخراج ویژگی

مهمترین مرحله در حل یک مسأله شناسایی الگو، مرحله انتخاب و استخراج ویژگی است. در این مقاله، از یک مجموعه کامل ویژگی بافت جهت طبقه‌بندی نواحی تصاویر در ساختار پیشنهادی [۱] استفاده شده است. این مجموعه شامل ویژگیهای Tamura (شامل زبری، جهت‌دار بودن و کنتراست)، همبستگی، انرژی، همگنی و آنتروپی مستخرج از ماتریس هم‌رخدادی [۱۶] و همچنین دو ویژگی مستخرج از زیرباندهای تبدیل موجک می‌باشد [۱].

۲-۲- تعریف ارزشمندی متن

در این مقاله، جهت انجام طبقه‌بندی معنایی تصاویر متنی شناسایی شده، از یک ساختار قابل تعریف ارزشمندی متن توسط کاربر استفاده می‌شود. بدین معنی که پس از انجام طبقه‌بندی نواحی استخراج شده از تصویر و مشخص شدن متنی یا عکسی بودن هر ناحیه از تصویر توسط الگوریتم طبقه‌بندی مبتنی بر ناحیه‌بندی، با مقایسه فازی این اطلاعات با الگوی از پیش تعریف شده ارزشمندی متن، میزان ارزشمند بودن متن مشخص می‌شود. بنابراین در تعریف الگوی ارزشمندی متن و فرآیند مقایسه فازی (به عبارت دیگر طبقه‌بندی معنایی تصویر) آن با یک تصویر، دو نکته نهفته است که در ادامه به آن پرداخته می‌شود. در این مقاله، تصاویر متنی مشابه به نامه‌های رسمی و اداری با ارزش تلقی می‌گردد. به عبارت دیگر، متن با ارزش باید حداقل شامل نواحی متنی و گرافیکی زیر باشد:

الف- یک ناحیه گرافیکی در بالای صفحه (در وسط یا سمت راست صفحه) که بیانگر علامت تجاری یا لوگوی یک شرکت یا سازمانی باشد.

۲-۳- طبقه‌بندی معنایی تصویر براساس الگوی ارزشمندی

مسأله‌ای که در طبقه‌بندی تصاویر متنی با ارزش و تصاویر متنی بی‌ارزش وجود دارد، بحث تعریف معیار یا به عبارتی الگوی با ارزش بودن متن موجود در تصویر است. همانگونه که مشخص است، تصمیم‌گیری نهایی در مورد محتوای یک تصویر ناشناخته، مسأله‌ای است که بصورت قابل توجهی به نظر و انتظار کاربر از سیستم طبقه‌بند دارد. بدین معنی که با توجه به تعریفی که کاربر از تصویر دلخواه یا غیردلخواه دارد، می‌توان تصمیم‌گیری در این مورد را انجام داد. در بخش قبل الگوی کلی ارزشمند بودن متن تشریح شد. طبق تعریف فوق، ویژگیهایی از الگوی ارزشمندی تعریف و استخراج می‌گردد تا به کمک آن بتوان یک شبکه عصبی پرسپترون چند لایه را آموزش داد. پس از انجام آموزش، از شبکه جهت تعیین مقدار با ارزش بودن تصویر ورودی استفاده می‌گردد. در ادامه با یک آستانه‌گذاری ساده بر روی این مقدار فازی می‌توان دسته با ارزش بودن (کلاس مطلوب) یا بی‌ارزش بودن (کلاس نامطلوب) تصاویر متنی را تعیین کرد. ویژگیهایی که از الگوی ارزشمندی استخراج می‌گردد به صورت زیر است:

الف- نسبت مساحت نواحی متنی و عکسی به کل تصویر.

ب- نسبت مساحت نواحی عکسی به نواحی متنی.

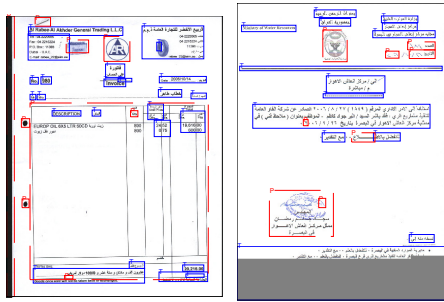
ج- مساحت هر ناحیه.

د- موقعیت هر ناحیه (مرکز جرم هر ناحیه).

ه- موقعیت نسبی مراکز جرم هر دو ناحیه (یعنی بالا، پایین، راست، چپ بودن).

و- نوع ناحیه (متنی یا گرافیکی بودن).

با بردارهای ویژگی شکل گرفته بصورت بالا، جهت آموزش یک شبکه عصبی استفاده می‌گردد. طبقه‌بند مذکور در سه لایه و با الگوریتم



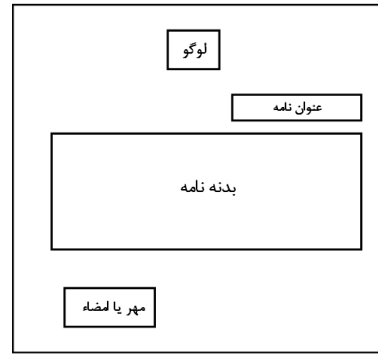
شکل(۴): عملکرد الگوریتم پیشنهادی ناحیه‌بندی و طبقه‌بندی بر روی نمونه تصاویر. نواحی متنی با مستطیلهای آبی رنگ با خطوط مستقیم و نواحی غیرمتنی با مستطیلهای قرمز رنگ با خطوط خط چین مشخص شده‌اند.

نتایج طبقه‌بندی نواحی مستخرج از تمام تصاویر پایگاه داده را نشان می‌دهد. در مجموع طبقه‌بندها ۱۹۸۴۲ ناحیه را به درستی شناسایی کرده‌اند. در حالیکه ۵۰۳ ناحیه را نتوانسته‌اند بدرستی شناسایی کنند. یعنی در مجموع صحت طبقه‌بندی برابر ۹۷/۵٪ بوده است و تنها ۲/۵٪ نواحی شناسایی نشده‌اند(جدول(۱)).

۳-۳- طبقه‌بندی معنایی تصاویر

جهت بررسی عملکرد طبقه‌بند معنایی پیشنهادی، مجموعه‌ای از تصاویر پایگاه داده با نرم‌افزار طبقه‌بند تصویر به دو گروه تصاویر متنی با ارزش(مطلوب) و غیر متنی(نامطلوب) طبقه‌بندی می‌شود (شکل(۵)). معیار مطلوب و نامطلوب بودن با توجه به تعریف الگوی ارزشمندی متن در نظر گرفته شده است. به این معنی که تصاویر حاوی نامه‌های رسمی و اداری به عنوان تصاویر با ارزش (تصاویر مطلوب) و بقیه تصاویر (شامل تصاویر windows و web، تصاویر خالص، روزنامه‌های اینترنتی و...) به عنوان تصاویر بی‌ارزش در نظر گرفته شده است. همانگونه که در شکل(۵) مشاهده می‌گردد، تصاویر خالص که دارای طیف رنگی وسیعی هستند در گروه تصاویر نامطلوب طبقه‌بندی شده‌اند. در مقابل، تصاویر با محتوا رنگی ضعیف در گروه تصاویر مطلوب قرار می‌گیرند. از طرف دیگر، تمام موارد شامل نامه‌های اداری و رسمی، گذرنامه و متنهای روزنامه‌ای به عنوان تصاویر با ارزش در نظر گرفته می‌شوند. تنها موردی که به نظر می‌رسد بی‌ارزش باشد، تصویر مربوط به یک مدار چاپی است که در گروه تصاویر با ارزش دسته‌بندی شده‌اند. تمام موارد شامل تصاویر رنگی، تصاویر فرمهای نرم‌افزارها و صفحات وبگاه‌های اینترنتی به عنوان تصاویر بی‌ارزش در نظر گرفته می‌شوند. در مجموع الگوریتم پیشنهادی، تصاویر متنی با ارزش را از بقیه تصاویر با صحت ۹۸/۸٪ طبقه‌بندی می‌کند.

در ادامه برای ارزیابی کامل‌تر الگوریتم پیشنهادی، از یک پایگاه داده با ۲۳۹۰ تصویر اینترنتی استفاده شده است که از نظر فراوانی تصاویر متنی و گرافیکی با دنیای واقعی مطابقت کامل دارد. جدول(۲) نتایج طبقه‌بندی ۲۳۹۰ تصویر اینترنتی را نشان می‌دهد. در این طبقه‌بندی از مجموع ۹ تصویر با ارزش تنها یک تصویر نوشتاری با کیفیت پایین



شکل(۳): الگوی نوعی یک متن با ارزش.

پس‌انتشار خطا آموزش می‌بیند تا در مورد نمونه‌های ناشناخته ورودی تصمیم‌گیری کند.

۳- نتایج آزمایشها

۳-۱- پایگاه داده

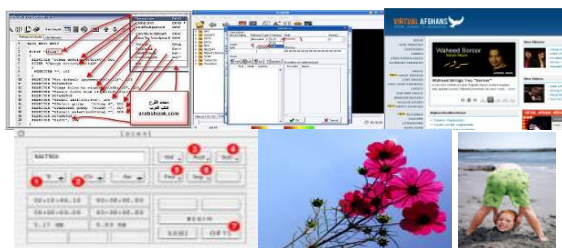
در این مقاله، جهت ارزیابی و آزمایش الگوریتم‌های پیشنهادی ناحیه‌بندی و طبقه‌بندی، از یک پایگاه داده تصاویر که از تصاویر موجود در شبکه اینترنت فراهم شده است، استفاده می‌شود. این مجموعه شامل ۹۷۷ سند تصویری است که به ترتیب شامل ۵۷۸ سند محض (یعنی تصویری که فقط شامل متن است)، ۸۱ تصویر محض (تصویر بدون متن) و ۳۱۸ سند شامل متن و تصویر (مانند روزنامه) می‌باشد. ۵۷۸ تصویر متنی خالص را می‌توان به اسناد دست‌نوشته، اسناد اداری، کارت‌های تبلیغاتی، روزنامه(چند ستونی)، اسناد تایپی، اسناد جدول‌بندی شده، فرم‌های ویندوزی و اینترنتی تقسیم نمود.

۳-۲- طبقه‌بندی نواحی تصویر

الگوریتم طبقه‌بندی مبتنی بر ناحیه‌بندی ارائه شده در [۱]، نواحی مختلف تصویر ورودی را شناسایی کرده و کلاس متنی یا گرافیکی بودن آن را مشخص می‌کند. در این قسمت، نتایج اعمال این الگوریتم بر روی چندین نمونه از تصاویر پایگاه داده نشان داده شده است. نمونه تصاویر بصورتی انتخاب شده است تا اینکه توانایی الگوریتم در حالتها و شرایط مختلف مورد ارزیابی قرار گیرد (شکل(۴)). در تصاویر ناحیه‌بندی شده شکل(۴) نواحی متنی با مستطیلهای آبی رنگ با خطوط مستقیم و نواحی غیرمتنی(نواحی عکسی) با مستطیلهای قرمز رنگ با خطوط خط‌چین نشان داده شده است. در تصاویر ناحیه‌بندی شده مشاهده می‌گردد که در اکثر موارد نواحی متنی بدرستی تشخیص داده شده‌اند و محل‌های امضاء، لوگوهای موجود در تصاویر و جداول به عنوان نواحی غیر متنی(نواحی شکلی) در نظر گرفته شده‌اند. در یکی از تصاویر(گل قرمز) نیز کل تصویر به عنوان ناحیه غیر متنی در نظر گرفته شده است و هیچ ناحیه متنی شناسایی نشده است. جدول (۱).

جدول (۱): صحت عملکرد طبقه‌بندی نواحی شناسایی شده.

نوع سند	تعداد موارد صحیح (TA)	تعداد موارد اشتباه (FA)	مجموع	صحت (%)
دست نوشته	۸۱۶	۲۳	۸۳۹	۹۷/۲۵
اداری	۱۶۸۰	۳۸	۱۷۱۸	۹۷/۷۸
کارت‌های تبلیغاتی	۱۴۲	۵	۱۴۷	۹۶/۶
روزنامه	۱۶۶۶	۳۰	۱۶۹۶	۹۸/۲۳
جدول بندی شده	۲۸۵۷	۷۵	۲۹۳۲	۹۷/۴۵
فرم‌های مبتنی بر web و windows	۴۰۴۶	۹۳	۴۱۳۹	۹۷/۷۵
شامل متن و تصویر باهم	۱۹۷۰	۷۵	۲۰۵۵	۹۵/۸۶
تصاویر محض	۲۲۶	۳۹	۲۶۵	۸۵/۲۸
اسناد دیگر	۶۴۳۹	۱۲۵	۶۵۶۴	۹۸/۰۹
جمع	۱۹۸۴۲	۵۰۳	۲۰۳۴۵	۹۷/۵۲



(ب)

شکل (۵): عملکرد الگوریتم طبقه‌بندی پیشنهادی بر روی یک مجموعه تصاویر آزمایشی. (الف) - تصاویر طبقه‌بندی شده با برچسب مطلوب و (ب) - تصاویر طبقه‌بندی شده با برچسب نامطلوب.

۴- بحث و جدل

دسترسی به اطلاعات موجود در تصاویر مسأله مهمی است که همواره مورد نظر محققان زمینه تحقیقاتی پردازش تصویر بوده است. در کاربرد تصاویر متنی، تنها ابزاری که برای نیل به این هدف مورد استفاده قرار می‌گیرد، روشهای تجزیه و تحلیل طرح‌بندی صفحات است. در این دسته روشها، تنها به نوع محتویات و موقعیت فیزیکی آنها پی می‌بریم در حالی که هیچ اطلاعاتی از مفاهیم و معانی نهفته در آنها در اختیار قرار نمی‌گیرد. در این مقاله به در نظر گرفتن این نکته، سعی شد به کمک ویژگیهای سطح پایینی همچون ویژگیهای بینایی تصویر مفاهیم سطح بالای موجود در آن تعریف شود. به عبارت دیگر یک تناظری بین مفاهیم سطح بالا و ویژگیهای سطح پایین برقرار گردد. موفقیت نسبی در برقراری این تناظر حداقل در دو مورد کمک شایانی به سیستم‌های جمع‌آوری اطلاعات می‌کند. اولاً، همانگونه که در بخش (۳-۳) آورده شد. اعمال این چنین الگوریتمی بر روی یک پایگاه داده بزرگ ۲۳۹۰ تصویری، باعث فیلتر شدن قسمت عمده‌ای از تصاویر بی‌ارزش موجود در مجاری اینترنت می‌شود. به عبارت دیگر، به جای اینکه کاربر ۲۳۹۰ تصویر را بصورت دستی مورد بررسی قرار دهد. با این الگوریتم تنها کافی کمتر از ۲٪ آن را یعنی ۴۵ تصویر را مورد بررسی دقیق‌تر قرار دهد. ثانیاً، ایجاد تناظر فوق وابستگی زیادی به نظرات کاربران دارد که این مسأله نکته مثبتی تلقی می‌گردد. چونکه

در دسته تصاویر بی ارزش طبقه‌بندی شده است. در مقابل نیز از مجموع ۲۳۸۱ تصویر بی‌ارزش تنها ۳۷ تصویر در دسته تصاویر با ارزش طبقه‌بندی شده‌است. به عبارت دیگر درصد پذیرش اشتباه (False Accept) بعنوان تصاویر با ارزش ۱/۵٪ می‌باشد. این نتایج نشان می‌دهد که کمتر از ۲٪ تصاویر اینترنتی برای بررسی توسط کارشناسان جدا می‌گردد.

جدول (۲): نتایج طبقه‌بندی تصاویر اینترنتی به دو گروه تصاویر با ارزش و بقیه تصاویر.

نوع تصویر	تعداد	نتایج طبقه بندی	
		متنی با ارزش	بقیه تصاویر
متنی با ارزش	۹	۸	۱
بقیه تصاویر	۲۳۸۱	۳۷	۲۳۴۴
مجموع	۲۳۹۰	۴۵	۲۳۴۵



(الف)

شبکه عصبی، تصاویر براساس محتوای متنی و عکسی، به دو گروه با ارزش و بی‌ارزش دسته‌بندی می‌شوند. الگوریتم پیشنهادی بر روی یک پایگاه داده تصاویر متنی و غیرمتنی که از تصاویر موجود در اینترنت فراهم شده است، مورد ارزیابی قرار گرفت. نتایج آزمایشها بیانگر کارایی روش پیشنهادی در طبقه‌بندی معنایی تصاویر براساس تعریف کاربر از با ارزش بودن و بی‌ارزش بودن تصاویر متنی است.

سیاسگزاری

این مقاله مستخرج از طرح پژوهشی مصوب در دانشگاه آزاد اسلامی واحد نجف‌آباد می‌باشد.

مراجع

- [۱] حسین پورقاسم، " تجزیه و تحلیل طرح‌بندی تصاویر متنی مبتنی بر طبقه‌بندی نواحی در یک ساختار سلسله مراتبی تصمیم‌گیری "، مجله الکترونیک- قدرت دانشگاه آزاد اسلامی واحد نجف‌آباد، شماره دوم، ص: ۳۷-۴۴، بهار ۱۳۸۸.
- [2] G. Nagy., "Twenty years of document image analysis in PAMI." *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 22, pp. 38-62, 2000.
- [3] N. Amamoto, S. Torigoe, and Y. Hirogaki, "Block Segmentation and Text Area Extraction of Vertically/Horizontally Written Document," *Proc. Second Int'l Conf. Document Analysis and Recognition*, pp. 739-742, Tsukuba, Japan, 1993.
- [4] M. Acharyya and M. K. Kundu, "Document Image Segmentation Using Wavelet Scale-Space Features," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 12, no. 12, pp. 1117-1127, 2002.
- [5] A. Busch, W. W. Boles, S. Sridharan, "Texture for script identification," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 27, no.11, pp.1720-1732, Nov. 2005.
- [6] S. Mao and T. Kanungo, "Empirical performance evaluation methodology and its application to page segmentation algorithms," *IEEE Trans. Pattern Analysis and Machine Intelligence*. Vol. 23, no.3, pp. 242-256, March 2001.
- [7] Y. Qiao, Z. Lu, C. Song, and S. Sun, "Document image segmentation using Gabor wavelet and kernel-based methods," *Systems and Control in Aerospace and Astronautics*, pp. 451-455, Jan 2006.
- [8] L. O'Gorman, "The Document Spectrum for Page Layout Analysis," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, pp. 1162-1173, 1993.
- [9] G. Nagy, S. Seth, and M. Viswanathan, "A prototype document image analysis system for technical journals", *Computer*, vol. 7, pp. 10-22, 1992.
- [10] A. Jain, and Y. Zhong, "Page Segmentation Using Texture Analysis," *Pattern Recognition*, vol. 29, pp. 743-770, 1996.
- [11] J. T. Laaksonen, J. M. Koskela, and E. Oja, "Picsom - a framework for content-based image database retrieval using self-organizing maps", *In Proceedings of SCIA'99, to be published*, Kangerlussuaq, Greenland, June 1999. URL: <http://www.cis.hut.fi/picsom/publications.html>.
- [12] Q. Ye, Q. Huang, W. Gao and D. Zhao, "Fast and robust text detection in images and video frames", *Image and Vision Computing*, Vol. 23, no. 6, pp. 565-576, June 2005.
- [13] H. Tamura, S. Mori, and T. Yamawaki, "Texture features corresponding to visual perception," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 8, No. 6, June 1978.
- [14] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Petkovic, P. Yanker, C. Faloutsos, and G. Taubin, "The QBIC project: Querying images by content using color, texture, and shape," *In Proceedings of the SPIE Conference on Storage and Retrieval for Image and Video Databases*, pp. 173-187, 1993.
- [15] A. Pentland, R.W. Picard, and S. Sclaroff, "Photobook: Content-based manipulation of image databases," *Int. Journal of Computer Vision*, vol. 18, no.3, pp. 233-254, June 1996.
- [16] R. M. Haralick, K. Shanmugan, and I. Dinstein, "Textural features for image classification", *IEEE Trans. on Systems, Man and Cybernetics*, vol.3, no.6, pp.610-621, 1973.
- [17] H. Li, D. Doermann, and O. Kia, "Automatic text detection and tracking in digital video", *IEEE Trans. Image Processing*, pp. 147-156, 2000.
- [18] R. Lienhart, A. Wernicke, "Localizing and segmenting text in images and videos", *IEEE Trans. Circuits and Systems for Video Technology*, pp. 256-268, 2002.

می‌توان با کمک گرفتن از نظرات کاربران در یک فرآیند بازخورد ربط جهت بهبود تناظر فوق استفاده کرد.

۵- نتیجه‌گیری

در این مقاله، جهت دستیابی به مفاهیم و معانی موجود در یک تصویر متنی، با استفاده از یک الگوی متنی با ارزش و نتایج حاصل از طبقه‌بندی نواحی تصویر، یک روش کارآمد جهت طبقه‌بندی تصاویر متنی با ارزش از بقیه تصاویر ارائه شد. در این ساختار، تصویر بوسیله الگوریتم ناحیه‌بندی دو مرحله‌ای، ناحیه‌بندی می‌شود. سپس متنی‌بودن و غیرمتنی (عکسی) بودن نواحی حاصل به کمک چندین طبقه‌بند در یک ساختار سلسله مراتبی طبقه‌بندی مشخص می‌گردد. در انتها، با استخراج ویژگی از الگوی متنی با ارزش و آموزش یک