Archive of SID

تجزیه و تحلیل طرح‌بندی تصاویر متنی مبتنی بر طبقه‌بندی نواحی در یک ساختار سلسله‌مراتبی تصمیم‌گیری

حسین پورقاسم
گروه برق - دانشگاه آزاد اسلامی واحد نجف‌آباد

خلاصه: تبدیل اسناد کاغذی به صورت نسخه الکترونیکی جهت ذخیره‌سازی، بازیابی و جستجو بر اساس محتوا مسأله مهمی در سیستم‌های خودکارسازی ادارای و سیستم‌هایی که وظیفه استخراج اطلاعات از مجاری اینترنت دارند، به شمار می‌رود. برای نیل به این هدف، ارائه سیستم‌هایی که بتواند محتوای تصاویر متنی[1] را تجزیه و تحلیل کنند، ضروری به نظر می‌رسد. در این مقاله، جهت تجزیه و تحلیل تصاویر متنی و دسترسی به محتوای آنها، یک ساختار سلسله‌مراتبی طبقه‌بندی مبتنی بر یک الگوریتم ناحیه‌بندی دو مرحله‌ای پیشنهاد شده است. در این ساختار، تصویر بوسیله الگوریتم پیشنهادی ناحیه‌بندی دومرحله‌ای، ناحیه‌بندی می‌شود. سپس متنی‌بودن و غیرمتنی (عکسی) بودن نواحی حاصل به کمک چندین طبقه‌بند در یک ساختار سلسله‌مراتبی طبقه‌بندی مشخص می‌گردد. الگوریتم ناحیه‌بندی پیشنهادی از دو مرحله ناحیه‌بندی مبتنی بر تبدیل موجک و آستانه‌گذاری استفاده می‌کند. از ویژگیهای بافت همچون ویژگی همبستگی، انرژی، همگنی و آنتروپی مستخرج از ماتریس همرخدادی و همچنین دو ویژگی مستخرج از زیرباندهای تبدیل موجک، جهت طبقه‌بندی و شناسایی برچسب نواحی شکل گرفته در مرحله ناحیه‌بندی استفاده می‌گردد. طبقه‌بند سلسه‌مراتبی از دو طبقه‌بند پرسپترون چندلایه[2] و یک طبقه‌بند ماشین بردار پشتیبان[3] تشکیل شده است. الگوریتم پیشنهادی بر روی یک پایگاه داده تصاویر متنی و غیرمتنی که از تصاویر موجود در اینترنت فراهم شده است، مورد ارزیابی قرار گرفته است. نتایج آزمایشها بیانگر کارایی روش پیشنهادی در ناحیه‌بندی تصاویر و طبقه‌بندی نواحی تصاویر است. الگوریتم پیشنهادی، صحت ۹۷/۵٪ را برای طبقه‌بندی نواحی تصاویر فراهم کرده است.

کلمات کلیدی: تجزیه و تحلیل طرح‌بندی صفحات، طبقه‌بندی، تصاویر متنی و غیرمتنی، ساختار سلسله‌مراتبی، ویژگیهای بافتی.

۱ - مقدمه

رشد فزاینده اسناد کاغذی دیجیتالی (اسناد کاغذی اسکن‌شده) در مجاری اینترنتی و سامانه‌های خودکارسازی ادارای، نیاز به بایگانی و دسترسی آنها جهت استخراج و داده‌کاوی اطلاعات را به یک نیاز مهم تبدیل کرده است. با توجه به زمان‌بر 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[4] جهت دسترسی به اطلاعات این اسناد امری اجتناب‌ناپذیر تلقی می‌شود. تجزیه و تحلیل اسناد اسکن‌شده، بخش وسیعی از اسناد، شامل روزنامه‌ها و مجلات، فاکس‌ها و نامه‌های تجاری، فرم‌های نرم‌افزاری، تصاویر پیوست‌شده به نامه‌های الکترونیکی، خطوط و نشانه‌های موسیقی، نقشه‌ها و رسم‌های فنی می‌شود [۱]. 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[5] 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

www.SID.ir

Archive of SID

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 [۱،۳].

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 داده‌اند (دسته روش سوم). در این تحقیقات ناحیه‌بندی معمولا به یکی از سه روش بالا به پایین،[۶] پایین به بالا[۷] و مبتنی بر بافت[۸] انجام شده است [۶،۷].

روش‌هایی که به صورت پایین به بالا عمل می‌کنند، از اجزای کوچک تصویر(پیکسل) استفاده کرده و با اتصال این اجزا به‌هم و تشکیل اجزای بزرگتر برای شناسایی قطعات اصلی موجود در تصویر استفاده می‌کنند. به عبارت دیگر این روشها، ناحیه‌بندی را از پیکسل شروع کرده و به ناحیه یا شئی می‌رسند. اشکال این روش در حساس بودن به اندازه حروف، وضوح سند اسکن‌شده و فاصله میان حروف است [۴]. از جمله این روش‌ها می‌توان به الگوریتم Docstrum اشاره کرد که براساس K- نزدیکترین همسایگی از اجزا صفحه عمل می‌کند [۸]. در روش‌هایی که به صورت بالا به پایین عمل می‌کنند، عملیات ناحیه‌بندی از بزرگترین جزء، یعنی تمام صفحه شروع شده تا به کوچکترین قطعه در صفحه می‌رسد. اشکال این روش در تشخیص زاویه و انحنای متن است. زیرا محدود به بلوکهای مستطیلی است [۴]. از جمله این روش‌ها می‌توان به الگوریتم بازگشتی X-Y Cut [۹] اشاره کرد. در هر مرحله مقادیر واریانس در جهت‌های افقی و عمودی محاسبه شده و با مقدار آستانه‌ای که از پیش تعریف شده است، مقایسه می‌گردد و در انتها نیز ناحیه‌بندی انجام می‌گیرد. روش‌هایی که بافت را برای ناحیه‌بندی بکارمی‌برند، از ویژگیهای بافت تصویر برای ناحیه‌بندی استفاده می‌کنند [۱۰]. در دو روش اول (پایین به بالا، بالا به پایین) نیاز به پیش اطلاعاتی درباره فاصله‌های افقی و عمودی، مجزا بودن متن‌ها، عکس‌ها، گرافیک‌ها و حتی فرض‌هایی درباره متن، سبک و اندازه می‌باشد. با توجه به اشکالات و موارد مطرح شده در مورد روشهای ناحیه‌بندی، در این مقاله از روش ناحیه‌بندی مبتنی بر بافت استفاده می‌شود. در بسیاری از روشهای ارائه‌شده قبلی در زمینه طبقه‌بندی، برای آنکه ویژگیها با اطلاعات محلی در اختیار قرار گیرند تا عملکرد طبقه‌بندی بهبود یابد، ویژگیها بصورت محلی استخراج می‌شوند [۱۱]. برای این منظور، تصویر به چندین ناحیه ثابت مثلا بلوکهای مستطیلی هم اندازه تقسیم شده و ویژگیها از هر بلوک استخراج می‌شوند. در این مقاله، برای استخراج ویژگیهای محلی و انجام طبقه‌بندی محتوایی، از یک نقشه طبقه‌بندی سلسله‌مراتبی مبتنی بر ناحیه‌بندی تصاویر سندی جهت تجزیه و تحلیل اسناد استفاده می‌شود. در این نقشه، از یک الگوریتم ناحیه‌بندی دومرحله‌ای مبتنی بر تبدیل موجک و آستانه‌گذاری استفاده می‌شود. در اولین مرحله از الگوریتم ناحیه‌بندی، از تبدیل موجک [۴-۵،۷،۱۲] برای ناحیه‌بندی استفاده می‌شود. سپس با استفاده از یک طبقه‌بند شبکه عصبی پرسپترون چندلایه (MLP)، برچسب نواحی از نظر متنی بودن یا گرافیکی(غیرمتنی) بودن مشخص می‌گردد و نواحی متنی شناسایی شده از تصویر کنار گذاشته می‌شوند. در مرحله دوم با اعمال یک ناحیه‌بندی ثانویه مبتنی بر آستانه‌گذاری بر روی باقی‌مانده تصویر حاصل از ناحیه‌بندی اولیه، تصویر بصورت کامل ناحیه‌بندی می‌شود. در انتها نیز برچسب نواحی ناحیه‌بندی شده توسط ناحیه‌بند دوم توسط دو طبقه‌بند (MLP و ماشین بردار پشتیبان (SVM)) مشخص می‌گردد. در هر سه طبقه‌بند، از ویژگیهای بافت جهت طبقه‌بندی نواحی استفاده می‌شود. ویژگیهایی همچون همبستگی، انرژی، همگنی و آنتروپی مستخرج از ماتریس همرخدادی و همچنین دو ویژگی از زیرباندهای تبدیل موجک استخراج می‌گردد.

در ادامه این مقاله، در بخش ۲ چارچوب پیشنهادی طبقه‌بندی سلسله‌مراتبی مبتنی بر ناحیه‌بندی معرفی می‌گردد. جزئیات ناحیه‌بندی دومرحله‌ای، استخراج ویژگی، ساختار طبقه‌بند سلسله‌مراتبی و طبقه‌بندی نهایی تصویر تشریح می‌شود. در بخش ۳ آزمایشهای انجام شده جهت ارزیابی الگوریتم پیشنهادی آورده شده است. در بخش ۴ مقایسه‌ای بین الگوریتم پیشنهادی با کارهای مشابه انجام شده قبلی صورت می‌گیرد. در انتها نیز در بخش ۵ نتیجه‌گیری از کار انجام شده در این مقاله ارائه می‌گردد.

۲- چارچوب پیشنهادی طبقه‌بندی سلسله‌مراتبی مبتنی بر ناحیه‌بندی

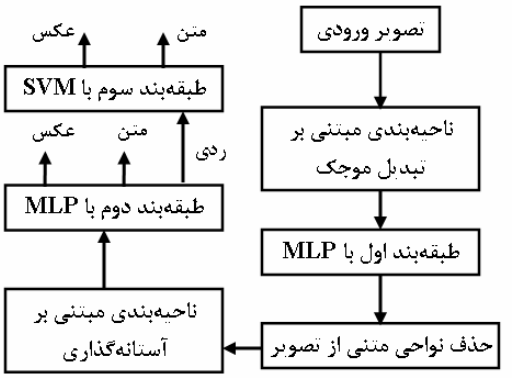
شکل (۱) روندنمای کلی چارچوب پیشنهادی برای طبقه‌بندی سلسله‌مراتبی مبتنی بر ناحیه‌بندی را نشان می‌دهد. در چارچوب پیشنهادی تصویر ورودی، طی دو مرحله ناحیه‌بندی، به یک سری نواحی با برچسبهای متنی و گرافیکی (غیرمتنی) ناحیه‌بندی می‌شود.

۲-۱- ناحیه‌بندی دومرحله‌ای

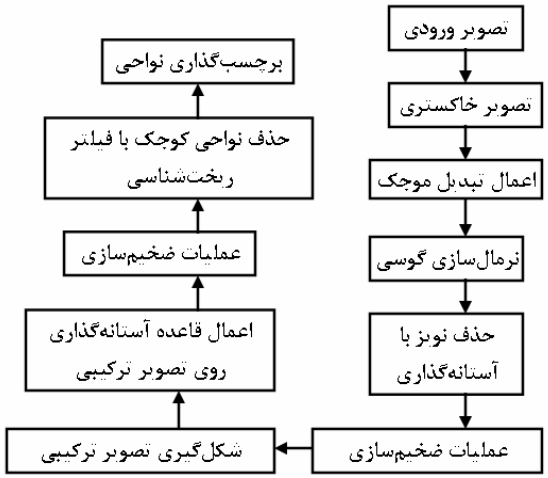
از یک الگوریتم ناحیه‌بندی دو مرحله‌ای برای ناحیه‌بندی تصویر و به دنبال آن طبقه‌بندی نواحی استفاده می‌شود. در مرحله اول این الگوریتم، تصویر ورودی با یک روش ناحیه‌بندی مبتنی بر تبدیل موجک ناحیه‌بندی می‌شود. سپس این نواحی با طبقه‌بند اول، طبقه‌بندی می‌شود. نواحی متنی شناسایی شده از تصویر ورودی حذف شده و نواحی گرافیکی در تصویر باقی می‌مانند تا دوباره و با دقت بیشتر توسط ناحیه‌بند دوم (ناحیه‌بند مبتنی بر آستانه‌گذاری) ناحیه‌بندی شده و به کمک دو طبقه‌بند (طبقه‌بند دوم و سوم)، متنی یا گرافیکی بودن نواحی مشخص می‌گردد.

www.SID.ir

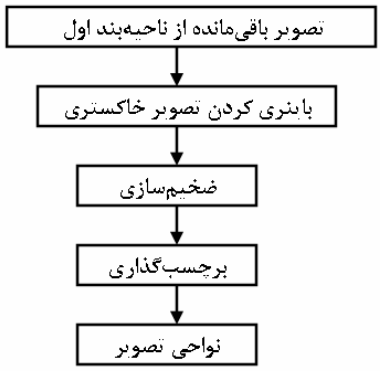
Archive of SID



شکل (۱): روندنمای طبقه‌بند سلسله‌مراتبی مبتنی بر ناحیه‌بندی.



شکل (۲): روندنمای الگوریتم ناحیه‌بندی مبتنی بر تبدیل موجک.



شکل (۳): روندنمای الگوریتم ناحیه‌بندی مبتنی بر آستانه‌گذاری.

۲-۱-۱- ناحیه‌بندی مبتنی بر تبدیل موجک

شکل (۲) روندنمای الگوریتم ناحیه‌بندی مبتنی بر تبدیل موجک را نشان می‌دهد. در این روندنما، تبدیل موجک دابیچی۲ در دو سطح بر روی تصویر خاکستری مربوط به تصویر ورودی اعمال می‌شود. عملیات نرمال‌سازی ضرایب تبدیل موجک (ضرایب افقی و عمودی در سطح اول H_1, V_1 و ضرایب افقی و عمودی در سطح دوم H_2, V_2) با یک تابع گوسی ($z = \frac{(x-\mu)}{\sigma}$) انجام می‌شود. در ادامه با اعمال آستانه $W_1 = \frac{\mu}{\sigma^2}$ بر روی ضرایب، ضرایب کوچک حذف می‌گردد. جهت تقویت ویژگیهای بافتی تصویر، از یک عملگر ریخت‌شناسی ضخیم‌سازی بصورت مربعی با اندازه ۳×۳ استفاده می‌گردد و ماتریس تصاویر $DilatedH_1$، $DilatedV_1$، $DilatedH_2$ و $DilatedV_2$ شکل می‌گیرد. سپس تصویر ترکیبی حاصل از مجموع زیرباندهای فوق بصورت زیر تعریف می‌شود:

$$HybridCoef = DilatedH_2 + DilatedV_2 \quad (1)$$

در ادامه جهت تعیین آستانه بر روی هر پیکسل از تصویر اصلی، تصویر ترکیبی شکل گرفته مطابق رابطه بالا (HybridCoef) و همچنین ماتریس ضرایب $DilatedH_1$ و $DilatedV_1$ با عملیات همسان‌سازی (نگاشت) هم‌اندازه می‌شوند. پس از همسازی‌سازی ماتریسهای فوق، با قاعدهٔ زیر مقادیر باینری تصویر ناحیه‌بندی شده شکل می‌گیرد: به ازای هر درایه از ماتریس HybridCoef، اگر مقادیر متناظر آن در هر دو ماتریس $DilatedH_1$ و $DilatedV_1$ بطور همزمان صفر باشند، مقدار آن با $Min(DilatedH_1, DilatedV_1)$ جایگزین می‌گردد در غیر این صورت مقدار قبلی خود را حفظ می‌کند. در ادامه، جهت اتصال نواحی کوچک نزدیک به یک ناحیه بزرگتر، بار دیگر از یک عملگر ریخت‌شناسی ضخیم‌سازی بصورت مربعی با اندازه ۴×۴ استفاده شده و بر روی ماتریس ترکیبی HybridCoef اعمال می‌گردد. معمولا، پس از اعمال عملگر ریخت‌شناسی ضخیم‌سازی بر روی تصویر ترکیبی، نواحی کوچک نویزی شکل می‌گیرد که با یک فیلتر ریخت‌شناسی با آستانه ۳۰۰ پیکسل حذف می‌شوند. در انتها، با برچسب‌گذاری نواحی شکل‌گرفته بر روی تصویر ترکیبی، نواحی جدا شده در تصویر اصلی مشخص و تعیین می‌گردند. با مجموعه ترفندهای فوق، مشکل چسبیدن نواحی متنی به گوشه‌های نواحی گرافیکی تصویر برطرف می‌گردد.

۲-۱-۲- ناحیه‌بندی مبتنی بر آستانه‌گذاری

پس از انجام ناحیه‌بندی اول یعنی ناحیه‌بندی مبتنی بر تبدیل موجک بر روی تصویر ورودی و حذف نواحی متنی شناسایی شده توسط این الگوریتم، الگوریتم ناحیه‌بندی دوم یعنی ناحیه‌بندی مبتنی بر آستانه‌گذاری بر روی باقی مانده تصویر اعمال می‌گردد. روندنمای الگوریتم ناحیه‌بندی مبتنی بر آستانه‌گذاری در شکل (۳) نشان داده شده است. در این الگوریتم، در ابتدا با پیدا کردن آستانه از روی هیستوگرام تصویر خاکستری، تصویر باینری می‌شود. سپس با عملگر ریخت‌شناسی ضخیم‌سازی به صورت مربعی با اندازه ۲۰×۲۰، نواحی کوچک نزدیک یکدیگر، بهم متصل می‌شوند.

www.SID.ir

Archive of SID

۲-۲- طبقه‌بندی سلسله‌مراتبی

مطابق با روندنمای شکل (۱)، در انتهای هر دو مرحله ناحیه‌بندی، نواحی شناسایی شده توسط سه طبقه‌بند دسته‌بندی می‌شوند. یک طبقه‌بند MLP بر روی خروجی‌های ناحیه‌بند اول و دو طبقه‌بند MLP و SVM بر روی خروجی‌های ناحیه‌بند دوم اعمال می‌شود. مراحل پیش‌پردازش و طبقه‌بندی نواحی در شکل (۴) نشان داده شده است. در هر یک از سه طبقه‌بند فوق، قبل از استخراج ویژگی از نواحی جهت طبقه‌بندی پیش‌پردازشهایی بر روی ناحیه انجام می‌شود. اولین پیش‌پردازش کنار گذاشتن نواحی کوچک با تغییرات سطوح خاکستری کم است. برای این منظور از فرکانس تغییرات سیاه به سفید و سفید به سیاه در راستاهای افقی و عمودی نواحی استفاده می‌شود و نواحی که فرکانس تغییرات کمتر از ۱۰ دارند، کنار گذاشته می‌شوند. در مرحله دوم از پیش‌پردازشها، جهت استخراج ویژگیهای مناسب از نواحی منظم، زاویه چرخش نواحی محاسبه می‌شود تا نواحی چرخیده شده اصلاح گردند. پس از اصلاح چرخش نواحی، ویژگیهای بافت از نواحی استخراج می‌گردند و توسط طبقه‌بند، طبقه‌بندی می‌شوند. طبقه‌بند اول که بر روی نواحی شناسایی شده توسط ناحیه‌بند اول عمل می‌کند، با اطمینان بالایی نواحی متنی را شناسایی می‌کند تا از تصویر حذف گردند. طبقه‌بند دوم که بر روی نواحی شناسایی شده توسط ناحیه‌بند دوم عمل می‌کند، نواحی را در سه گروه متنی، گرافیکی و ردی دسته‌بندی می‌کند. نواحی ردی توسط طبقه‌بند سوم به دو گروه متنی و گرافیکی طبقه‌بندی می‌شود.

۲-۲-۱- 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 استفاده شده است. این مجموعه شامل همبستگی، انرژی، همگنی و آنتروپی مستخرج از ماتریس همرخدادی [۱۳] و همچنین دو ویژگی مستخرج از زیرباندهای تبدیل موجک می‌باشد. در ادامه جزئیات استخراج این ویژگی تبدیل موجک آورده می‌شود.

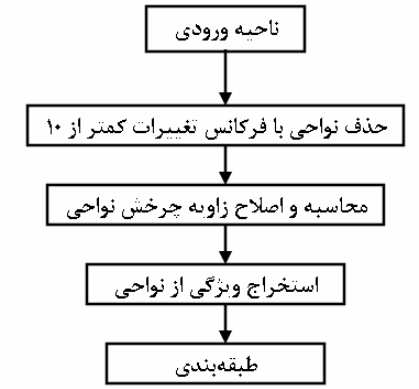
الف- ویژگیهای تبدیل موجک

در این مقاله، علاوه بر ویژگیهای تامورا و همرخدادی، از ویژگیهای مستخرج از تبدیل موجک نیز استفاده می‌شود. برای استخراج این ویژگیها، پس از تجزیه تصویر به زیرباندهای تبدیل موجک، مقادیر میانگین و انحراف معیار هر زیرباند محاسبه می‌شود. با توجه به اینکه تبدیل موجک در سه سطح بر روی تصویر اعمال می‌گردد در آن صورت با توجه به اینکه ۳×۳ دسته ضریب جزئیات در سه راستای افقی، عمودی و قطری وجود دارد بنابراین ۳×۳×۲ ویژگی شکل می‌گیرد. در این روش از دسته ضرایب تقریب برای استخراج ویژگی استفاده نمی‌شود.

$$\text{False Negative Rejection} = \frac{\text{Number of negative rejection}}{\text{Number of rejection}} \qquad (۱)$$

۲-۳- طبقه‌بندی نواحی تصاویر

در انتهای هر مرحله ناحیه‌بندی با استخراج ویژگیهای تشریح‌شده،



شکل (۴): مراحل پیش‌پردازش و طبقه‌بندی نواحی.

اقدام به انجام عمل طبقه‌بندی می‌گردد. برای انجام عمل طبقه‌بندی از دو طبقه‌بند MLP و SVM استفاده شده است. در هر یک از دو طبقه‌بند فوق، طبقه‌بند با ویژگیهای مستخرج از نواحی شناسایی شده مجموعه آموزشی، آموزش دیده تا پارامترهای بهینه هر یک از آنها استخراج گردد تا در مرحله آزمایش مورد استفاده قرار گیرد.

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 جدول‌بندی‌شده، فرم‌های ویندوزی و اینترنتی تقسیم نمود. نمونه‌ای از انواع تصاویر پایگاه داده در شکل (۵) نشان داده شده است.

۳-۲- معیارهای ارزیابی الگوریتم طبقه‌بندی

معمولا برای ارزیابی عملکرد روشهای طبقه‌بندی تصاویر از معیارهای Precision-Recall استفاده می‌شود که بصورت زیر تعریف می‌شود [۱۲].

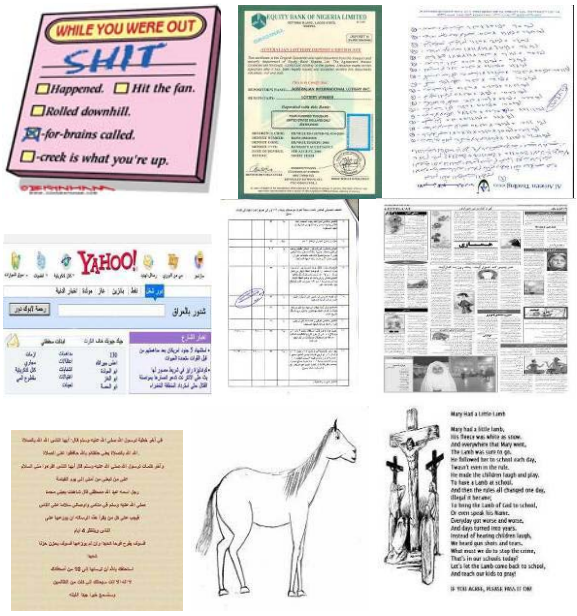
$$\text{Precision - Recall} = \frac{\text{Number of correctly detected}}{\text{Number of detected}} \qquad (۲)$$

$$\text{False Accept} = \frac{\text{Number of falsely detected}}{\text{Number of detected}} \qquad (۳)$$

$$\text{True Positive Rejection} = \frac{\text{Number of true positive rejection}}{\text{Number of rejection}} \qquad (۴)$$

www.SID.ir

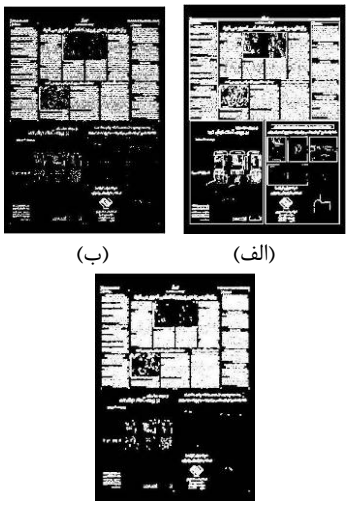
Archive of SID



شکل (۵): نمونه‌ای از انواع مختلف تصاویر پایگاه داده.

(ب) (الف)

شکل (۶): تصویر نمونه روزنامه‌ای جهت بررسی نتایج میانی الگوریتم ناحیه‌بندی دومرحله‌ای. (الف)- تصویر نمونه (ب)- تصویر خاکستری معکوس‌شده.



(ب) (الف)

(ج)

شکل (۷): نتایج اعمال ناحیه‌بند مرحله اول بر روی تصویر. (الف)- تصویر ترکیبی حاصل از یکجا کردن زیرباندهای تبدیل موجک (ب)- تصویر ترکیبی پس از اعمال قاعده آستانه‌گذاری (ج)- تصویر نهایی ناحیه‌بندی شده توسط ناحیه‌بند اول.

۳-۳- ناحیه‌بندی دومرحله‌ای

جهت بررسی نتایج میانی مراحل مختلف الگوریتم ناحیه‌بندی دو مرحله‌ای پیشنهادی از یک تصویر نمونه روزنامه‌ای استفاده می‌شود (شکل (۶- الف)). در این تصویر نمونه، نواحی متنی و عکسی بسیاری در نزدیک یکدیگر و در برخی موارد نیز ناحیه متنی درون یک ناحیه عکسی قرار گرفته است. با توجه به اینکه پس زمینه سفید رنگ است و پردازشها براساس اشیاء سفید رنگ انجام می‌گردد. تصویر خاکستری شده معکوس می‌گردد و پردازشها بر روی آن انجام می‌شود. نتایج میانی اعمال ناحیه‌بند اول بر روی تصویر نمونه در شکل (۷) نشان داده شده است. تصویر ترکیبی حاصل از یکجا کردن زیرباندهای تبدیل موجک، تصویر ترکیبی پس از اعمال قاعده آستانه‌گذاری و تصویر نهایی ناحیه‌بندی‌شده توسط ناحیه‌بند اول در شکل (۷) نشان داده شده است. همانگونه که مشاهده می‌کنید که نواحی متنی تصویر در این مرحله از ناحیه‌بندی با دقت خوبی از نواحی عکسی و بقیه نواحی متنی جدا می‌شوند. پس از حذف نواحی متنی شناسایی شده توسط طبقه‌بند اول از تصویر، ناحیه‌بند دوم بر روی باقی‌مانده تصویر اعمال می‌گردد. نتایج ناحیه‌بندی دوم بر روی باقی‌مانده تصویر حاصل از ناحیه‌بندی اول در شکل (۸) نشان داده شده است. همانگونه که در شکل (۸) مشاهده می‌گردد، همانطور که در مرحله اول ناحیه‌بندی نواحی متنی به خوبی جدا و طبقه‌بندی می‌شوند. در ناحیه‌بند دوم نواحی عکسی با دقت بالایی از بقیه تصویر جدا می‌شوند. در انتها، نواحی شناسایی شده توسط الگوریتم ناحیه‌بندی دومرحله‌ای (متنی و عکسی) بر روی تصویر اصلی با مستطیل‌های قرمز رنگ مشخص شده است (شکل (۹)). همانگونه که در نتایج شکل (۹) مشاهده می‌گردد، در برخی از نواحی بزرگ، نواحی متنی و گرافیکی کوچکتر نیز شناسایی می‌شود که به صورت مجزا به طبقه‌بندی داده می‌شود تا نوع آن ناحیه مشخص گردد.

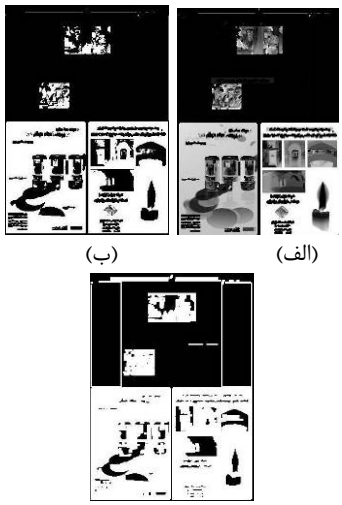
۳-۴- ارزیابی الگوریتم طبقه‌بندی نواحی مستخرج از ناحیه‌بندی

برای ارزیابی الگوریتم پیشنهادی ناحیه‌بندی و همچنین طبقه‌بندی، الگوریتم بر روی چندین نمونه از تصاویر پایگاه داده آزمایش می‌شود. نمونه تصاویر به‌صورتی انتخاب شده است تا اینکه توانایی الگوریتم در حالتها و شرایط مختلف مورد ارزیابی قرار گیرد. در شکل (۱۰) چندین نمونه تصویر به همراه نواحی مشخص شده با الگوریتم ناحیه‌بندی و برچسب آنها که با طبقه‌بندی تعیین شده است، نشان داده شده است. در تصاویر ناحیه‌بندی شده شکل (۱۰) نواحی متنی با مستطیلهای آبی رنگ با خطوط ممتد 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 غیرمتنی (نواحی شکلی) در نظر

www.SID.ir

Archive of SID

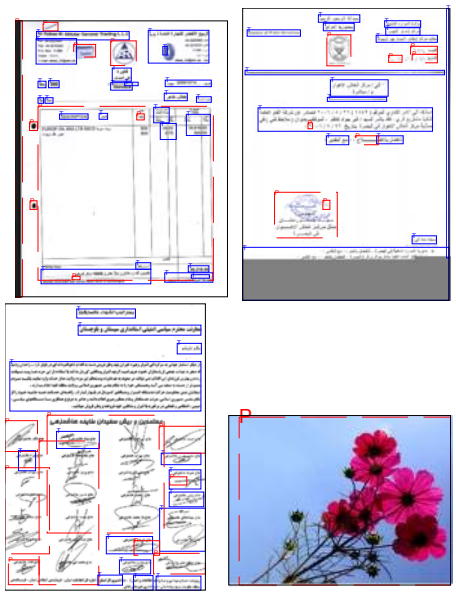
گرفته شده‌اند. در یکی از تصاویر (گل قرمز) نیز کل تصویر به عنوان ناحیه غیرمتنی در نظر گرفته شده است و هیچ ناحیه متنی شناسایی نشده است.



(الف) (ب) (ج)

شکل (۸): نتایج ناحیه‌بندی دوم برروی باقی‌مانده تصویر حاصل از ناحیه‌بندی اول. (الف)- حذف نواحی متنی از تصویر ورودی، (ب)- تصویر باینری شده با استفاده از هیستوگرام رنگ، (ج)- تصویر پس از عمل ضخیم‌سازی.

شکل (۹): نواحی شناسایی شده توسط ناحیه‌بندی دومرحله‌ای.



شکل (۱۰): عملکرد الگوریتم پیشنهادی ناحیه‌بندی و طبقه‌بندی بر روی نمونه تصاویر. نواحی متنی با مستطیلهای آبی رنگ با خطوط ممتد و نواحی غیرمتنی با مستطیلهای قرمز رنگ با خطوط خط‌چین مشخص شده‌اند.

۳-۵- طبقه‌بندی نواحی حاصل از ناحیه‌بندی

در این قسمت عملکرد طبقه‌بند سلسله‌مراتبی در طبقه‌بندی نواحی حاصل از الگوریتم ناحیه‌بندی دومرحله‌ای مورد بررسی قرار می‌گیرد. نتایج طبقه‌بندی نواحی شکلی، متنی و نواحی ردشده به صورت مجزا ارائه می‌گردد. نتایج طبقه‌بندی نواحی متنی، تصویری و نواحی رد شده برای نمونه تصاویر مختلف پایگاه داده، به صورت مجزا در جداول (۱) و (۲) نشان داده شده است.

طبقه‌بندها در مجموع ۹۷/۵٪ ناحیه را به درستی تشخیص داده‌اند (جدول (۱)). در جدول (۲)، میزان TPR و FNR به ترتیب برابر با ۷۰/۴٪ و ۲۹/۶٪ می‌باشد. 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. البته به دلیل وجود ناحیه‌بند دوم، مقدار FNR تاثیر چندانی در کارایی روش پیشنهادی ندارد. زیرا نواحی رد شده در مرحله اول توسط ناحیه‌بند دوم مجددا ناحیه‌بندی می‌شوند که باعث کاهش اثر نواحی ردشده در کارایی روش پیشنهادی می‌شود.

۴- بحث

جدول (۳)، مقایسه‌ای بین عملکرد الگوریتم پیشنهادی با الگوریتم‌های ارائه شده در مقالات ارائه می‌کند. نتایج حاصل از الگوریتم پیشنهادی بهبود ۳/۳٪ را نسبت به بهترین نتایج الگوریتم‌های قبلی را نشان می‌دهد. این بهبود قابل توجه را می‌توان در استفاده مناسب و به‌جا از ساختار طبقه‌بندی سلسله‌مراتبی به همراه ویژگیهای خوب مورد استفاده قرار داد. علاوه بر این استفاده از یک الگوریتم ناحیه‌بندی دومرحله‌ای به همراه طبقه‌بندی نواحی و سپس تصمیم‌گیری در مورد تصویر بر اساس الگوی نواحی متنی و عکسی در این بهبود تأثیر زیادی داشته است. در اکثر الگوریتم‌هایی که تاکنون استفاده شده است، بیشتر جهت تشخیص متن می‌باشند و کمتر به بررسی نواحی غیرمتنی (تصویری) پرداخته شده است. اما، ما در این مقاله الگوریتمی جهت شناسایی و طبقه‌بندی نواحی تصویر ارائه کردیم. در کارهای قبلی برآوردی از هزینه محاسباتی و زمانی الگوریتمها گزارش نشده است. ولی به نظر می‌رسد که از نظر پیچیدگی محاسباتی، الگوریتم پیشنهادی دارای پیچیدگی بیشتری نسبت به الگوریتم‌های قبلی می‌باشد. دلیل این افزایش را می‌توان در دومرحله‌ای بودن ناحیه‌بندی و همچنین سلسله‌مراتبی بودن طبقه‌بندها جستجو کرد. اما در مقابل با استفاده از ویژگی‌های متناسب با هر دسته از تصاویر و چندین طبقه‌بند مختلف در یک ساختار سلسله‌مراتبی عملکرد طبقه‌بندی بصورت قابل توجهی بهبود می‌یابد.

www.SID.ir

Archive of SID

جدول (۱): صحت عملکرد طبقه‌بندی نواحی شناسایی‌شده در ۹ دسته سند.

تعداد سند	تعداد مواردصحیح (TA)	تعداد موارد اشتباه (FA)	جمع کل	بازیابی-دقت (٪)
۹۷۷	۱۹۸۴۲	۵۰۳	۲۰۳۴۵	۹۷/۵۲

جدول (۲): نواحی ردشده توسط طبقه‌بندها.

تعداد سند	تعداد موارد اشتباه (FR)	تعداد موارد صحیح (FA)	جمع کل	(٪) FR	(٪) FA
۹۷۷	۳۳۶	۷۹۸	۱۱۳۴	۲۹/۶	۷۰/۴

جدول (۳): مقایسه الگوریتم پیشنهادی با الگوریتم‌های قبلی.

الگوریتم	طبقه‌بند مورد استفاده	ویژگی مورد استفاده	صحت طبقه‌بندی (٪)
الگوریتم[۱۲]	SVM	همرخدادی و موجک	۹۴/۲
الگوریتم[۱۴]	SVM	موجک	۹۱/۴
الگوریتم[۱۵]	SVM	گرادیان	۹۴/۳
الگوریتم پیشنهادی	SVM & Neural network	موجک و همرخدادی	۹۷/۵

۵- نتیجه‌گیری

در این مقاله، جهت تجزیه و تحلیل تصاویر متنی و دسترسی به محتوای آنها، یک ساختار سلسله‌مراتبی طبقه‌بندی مبتنی بر یک الگوریتم ناحیه‌بندی دو مرحله‌ای پیشنهاد شد. در این ساختار، تصویر بوسیله الگوریتم پیشنهادی ناحیه‌بندی دو مرحله‌ای، ناحیه‌بندی می‌شود. سپس متنی‌بودن و غیرمتنی (عکسی) بودن نواحی حاصل به کمک چندین طبقه‌بند در یک ساختار سلسله‌مراتبی طبقه‌بندی مشخص می‌گردد. در انتها، با آموزش الگوی نواحی متنی و غیرمتنی در تصاویر، تصاویر براساس محتوای متنی و عکسی، به دو گروه با ارزش و بی‌ارزش دسته‌بندی می‌شود. الگوریتم ناحیه‌بندی پیشنهادی از دو مرحله ناحیه‌بندی مبتنی بر تبدیل موجک و آستانه‌گذاری استفاده می‌کند. از ویژگیهای بافت همچون همبستگی، انرژی، همگنی و آنتروپی مستخرج از ماتریس همرخدادی و همچنین دو ویژگی مستخرج از زیرباندهای تبدیل موجک، جهت طبقه‌بندی و شناسایی برچسب نواحی شکل گرفته توسط الگوریتم ناحیه‌بندی استفاده می‌شود. طبقه‌بند سلسه‌مراتبی از دو طبقه‌بند پرسپترون چندلایه و یک طبقه‌بند ماشین‌بردار پشتیبان استفاده می‌کند. 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 آزمایشات بیانگر کارایی روش پیشنهادی در ناحیه‌بندی تصاویر و نهایتاً طبقه‌بندی نواحی تصاویر است.

سپاسگزاری

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

پی‌نوشت :

1- Document Image
2- Multi layer Perceptron
3- Support Vector Machine
4- Document Image Analysis
5- Layout
6- Top-Down
7- Bottom-Up
8- Texture

www.SID.ir

Archive of SID

مراجع

[1] A.K.Jain, B.Yu, "Document representation and its application to page decomposition", IEEE Trans. Pattern Analysis and Mac. Intel., Vol.20, pp.294–308, 1998.

[2] G.Nagy, "Twenty years of document image analysis in PAMI", IEEE Trans. Pattern Analysis and Machine Intelli., Vol. 22, pp. 38-62, 2000.

[3] N. Amamoto, S. Torigoe, Y. Hirogaki, "Block segmentation and text area extraction of vertically/horizontally written document," ICDAR, pp.739-742, Tsukuba, Japan, 1993.

[4] M.Acharyya, M.K.Kundu, "Document image segmentation using wavelet scale-space features", IEEE Trans. Cir. and Sys. for Vid. Tech., Vol.12, No.12, pp.1117-1127, 2002.

[5] A.Busch, W.W.Boles, S.Sridharan, "Texture for script identification", IEEE Trans. Patt. Ana. and Mac. Intelli., Vol. 27, No.11, pp.1720-1732, Nov. 2005.

[6] S.Mao, T.Kanungo, "Empirical performance evaluation methodology and its application to page segmentation algorithms", IEEE Trans. Patt. Anal. and Mac. Intel., Vol.23, No.3, pp.242-256, March 2001.

[7] Y.Qiao, Z.Lu, C.Song, S.Sun, "Document image segmentation using Gabor wavelet and kernel-based methods," Syst. and Control in Aeros. and Astr., pp. 451-455, Jan 2006.

[8] L. O'Gorman, "The document spectrum for page layout analysis," IEEE Trans. Patt. Ana. and Mac. Intell., 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," Patte. Reco., Vol. 29, pp. 743-770, 1996.

[11] J.T.Laaksonen, J.M.Koskela, E.Oja, "Picsom - A framework for content-based image database retrieval using self-organizing maps", In Proceedings of SCIA, to be published, Kangerlussuaq, Greenland, June 1999. URL: http://www.cis.hut.fi/picsom/publications.html.

[12] Q.Ye, Q.Huang, W.Gao, D.Zhao, "Fast and robust text detection in images and video frames", Ima. and Vis. Comp., Vol.23, No.6, pp.565-576, June 2005.

[13] R.M.Haralick, K.Shanmugan, I.Dinstein, "Textural features for image classification", IEEE Trans. on Syst., Man and Cybe., Vol.3, No.6, pp.610-621, 1973.

[14] H.Li, D.Doermann, O.Kia, "Automatic text detection and tracking in digital video", IEEE Trans.Ima. Proc., pp.147-156, 2000.

[15] R.Lienhart, A.Wernicke, "Localizing and segmenting text in images and videos", IEEE Trans. Circ. and Syst. for Vid. Tech., pp.256-268, 2002.

رزومه

- حسین پورقاسم استادیار گروه مخابرات دانشکده برق دانشگاه آزاد اسلامی واحد نجف‌آباد، دکتری تخصصی برق با گرایش مهندسی پزشکی را در سال ۱۳۸۷ از دانشگاه تربیت مدرس اخذ کرده است. وی مدرک کارشناسی و کارشناسی ارشد خود را در همین رشته در سالهای ۱۳۸۱ و ۱۳۸۳ دریافت کرده است. نامبرده در تمامی این دوره‌ها دانشجوی ممتاز بوده است. زمینه‌های تحقیقاتی مورد علاقه وی شامل زمینه‌های مختلف پردازش تصویر مانند نمایه‌گذاری و بازیابی تصاویر، بیومتریک، ردیابی اشیاء، مخفی کردن اطلاعات، شناسایی الگو و شبکه‌های عصبی مصنوعی می‌باشد و در این زمینه‌ها تاکنون بیش از ۲۰ مقاله در مجلات و کنفرانسهای خارجی و داخلی به چاپ رسانیده است.

www.SID.ir