

تجزیه و تحلیل طرح بندی تصاویر متنی مبتنی بر طبقه بندی نواحی در یک ساختار سلسله مراتبی تصمیم گیری

حسین پورقاسم

استادیار گروه برق - دانشکده برق - دانشگاه آزاد اسلامی واحد نجف آباد

تاریخ پذیرش: بهار ۱۳۸۹

تاریخ دریافت: تابستان ۱۳۸۸

خلاصه: تبدیل اسناد کاغذی به صورت نسخه الکترونیکی جهت ذخیره سازی، بازیابی و جستجو بر اساس محتوا مسأله مهمی در سیستم های خودکارسازی اداری و سیستم هایی که وظیفه استخراج اطلاعات از مجاری اینترنت دارند، به شمار می رود. برای نیل به این هدف، ارائه سیستم هایی که بتواند محتوای تصاویر متنی^۱ را تجزیه و تحلیل کنند، ضروری به نظر می رسد. در این مقاله، جهت تجزیه و تحلیل تصاویر متنی و دسترسی به محتوای آنها، یک ساختار سلسله مراتبی طبقه بندی مبتنی بر یک الگوریتم ناحیه بندی دو مرحله ای پیشنهاد شده است. در این ساختار، تصویر به وسیله الگوریتم پیشنهادی ناحیه بندی دو مرحله ای، ناحیه بندی می شود. سپس متنی بودن و غیرمتنی (عکسی) بودن نواحی حاصل به کمک چندین طبقه بندی در یک ساختار سلسله مراتبی طبقه بندی مشخص می گردد. الگوریتم ناحیه بندی پیشنهادی از دو مرحله ناحیه بندی مبتنی بر تبدیل موجک و آستانه گذاری استفاده می کند. از ویژگیهای بافت همچون ویژگی همبستگی، انرژی، همگنی و آنتروپی مستخرج از ماتریس همخدادی و همچنین دو ویژگی مستخرج از زیرباند های تبدیل موجک، جهت طبقه بندی و شناسایی برچسب نواحی شکل گرفته در مرحله ناحیه بندی استفاده می گردد. طبقه بندی سلسله مراتبی از دو طبقه بندی پرسپترون چندلایه^۲ و یک طبقه بندی ماشین بردار پشتیبان^۳ 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 ناحیه بندی تصاویر و طبقه بندی نواحی تصاویر است. الگوریتم پیشنهادی، صحت 97.5٪ را برای طبقه بندی نواحی تصاویر فراهم کرده است.

کلمات کلیدی: تجزیه و تحلیل طرح بندی صفحات، طبقه بندی، تصاویر متنی و غیرمتنی، ساختار سلسله مراتبی، ویژگیهای بافتی.

۱ - مقدمه

رشد فزاینده اسناد کاغذی دیجیتالی (اسناد کاغذی اسکن شده) در مجاری اینترنتی و سامانه های خودکارسازی اداری، نیاز به بایگانی و دسترسی آنها جهت استخراج و داده کاوی اطلاعات را به یک نیاز مهم تبدیل کرده است. با توجه به زمان بر بودن این فرآیند به صورت دستی (توسط کاربر)، ارائه ابزار خودکاری که بتواند این نیاز را مرتفع کند، ضروری به نظر می رسد. بنابراین در یک سامانه خودکار جستجو و بازیابی اطلاعات متنی، تجزیه و تحلیل اسناد متنی^۴ جهت دسترسی به اطلاعات این اسناد امری اجتناب ناپذیر تلقی می شود. تجزیه و تحلیل اسناد اسکن شده، بخش وسیعی از اسناد، شامل روزنامه ها و مجلات، فاکس ها و نامه های تجاری، فرم های نرم افزاری، تصاویر پیوست شده به نامه های الکترونیکی، خطوط و نشانه های موسیقی، نقشه ها و رسم های فنی می شود [۱]. اسناد به سه روش عمده تجزیه و تحلیل می شوند [۲]:

۱- تجزیه و تحلیل بر اساس محتویات سند: در این دسته روش صرف نظر از اینکه چطور اسناد ایجاد شده اند، براساس محتویاتشان به دو دسته تقسیم می شوند: طبیعی (مانند عکس صورت و تصاویر ماهواره ای) و نمادین (مانند کارهای چاپی، فرم های اداری و طراحی های مهندسی) [۲].

۲- تجزیه و تحلیل بر اساس لایه بندی^۵ سند: در این دسته روش، لایه بندی سند مورد بررسی قرار می گیرد. این دسته روش نیز به نوبه خود به دو دسته تقسیم می شود: 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 دسته بندی می کند. در مقابل، تجزیه و تحلیل منطقی، به نوع صفحه بندی می پردازد و برچسب های موجود در صفحه مانند عنوان، علامت تجاری، پاورقی، حروف بزرگ و کوچک و ... را مشخص می کند [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 معمولاً 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 به هم 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 اسکن شده 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 به کار می‌برند، 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 به صورت محلی استخراج می‌شوند [۱۱].

برای این منظور، تصویر به چندین ناحیه ثابت مثلاً 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 [۴-۷، ۱۲، ۱۳] 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 به صورت 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 آنترپی مستخرج از ماتریس هم‌رخدادی و همچنین دو ویژگی از زیرباندای 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

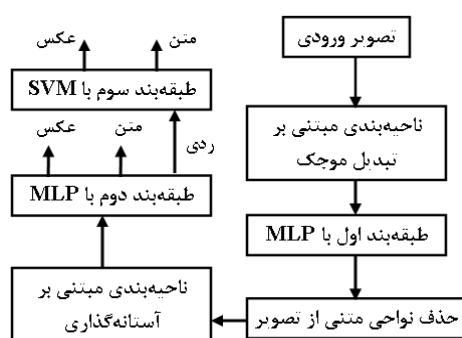
شکل (۱) روندنمای کلی چارچوب پیشنهادی برای طبقه‌بندی سلسله‌مراتبی مبتنی بر ناحیه‌بندی را نشان می‌دهد. در چارچوب پیشنهادی تصویر ورودی، طی دو مرحله ناحیه‌بندی، به یک سری نواحی با برچسبهای متنی و گرافیکی (غیرمتنی) ناحیه‌بندی می‌شود.

۲-۱- ناحیه‌بندی دومرحله‌ای

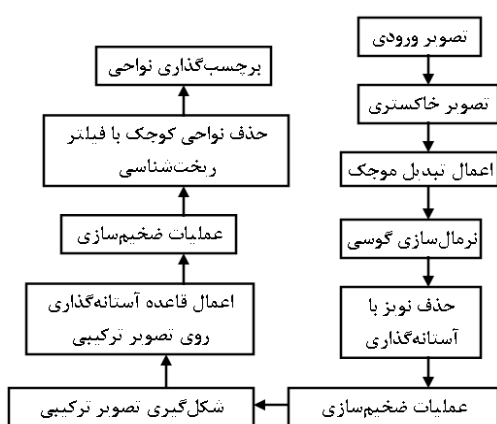
از یک الگوریتم ناحیه‌بندی دو مرحله‌ای برای ناحیه‌بندی تصویر و به دنبال آن طبقه‌بندی نواحی استفاده می‌شود. در مرحله اول این الگوریتم، تصویر ورودی با یک روش ناحیه‌بندی مبتنی بر تبدیل موجک ناحیه‌بندی می‌شود. سپس این نواحی با طبقه‌بند اول، طبقه‌بندی می‌شود. نواحی متنی شناسایی شده از تصویر ورودی حذف شده و 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.

۲-۱-۱- ناحیه‌بندی مبتنی بر تبدیل موجک

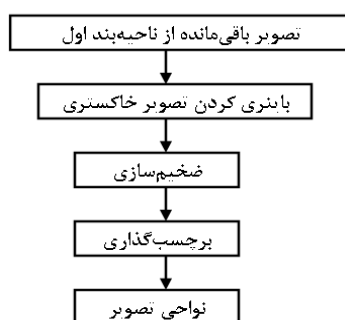
شکل (۲) روندنمای الگوریتم ناحیه‌بندی مبتنی بر تبدیل موجک را نشان می‌دهد. در این روندنما، تبدیل موجک دایچی ۲ در دو سطح بر روی تصویر خاکستری مربوط به تصویر ورودی اعمال می‌شود. عملیات نرمال‌سازی ضرایب تبدیل موجک (ضرایب افقی و عمودی در سطح اول V_1, H_1 و ضرایب افقی و عمودی در سطح دوم V_2, H_2) با یک



شکل (۱): روند نمای طبقه بندی سلسله مراتبی مبتنی بر ناحیه بندی
Fig. (1): flowchart of the classification based segmentation



شکل (۲): روند نمای الگوریتم ناحیه بندی مبتنی بر تبدیل موجک
Fig. (2): segmentation algorithm flowchart based on wavelet transform



شکل (۳): روند نمای الگوریتم ناحیه بندی مبتنی بر آستانه گذاری.
Fig. (3): segmentation algorithm flowchart based on thresholding

۲-۲- طبقه بندی سلسله مراتبی

مطابق با روند نمای شکل (۱)، در انتهای هر دو مرحله ناحیه بندی، نواحی شناسایی شده توسط سه طبقه بندی دسته بندی می شوند. یک طبقه بندی MLP بر روی خروجی های ناحیه بندی اول و دو طبقه بندی SVM و بر روی خروجی های ناحیه بندی دوم اعمال می شود. مراحل پیش پردازش و طبقه بندی نواحی در شکل (۴) نشان داده شده است. در هر یک از سه طبقه بندی فوق، قبل از استخراج ویژگی از نواحی جهت

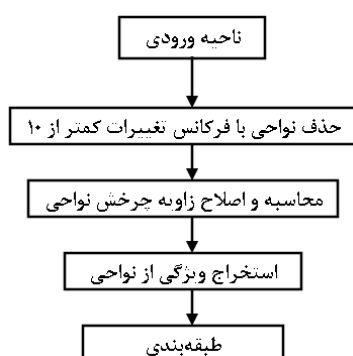
تابع گوسی $(z = \frac{x-\mu}{\sigma})$ انجام می شود. در ادامه با اعمال آستانه $W_1 = \frac{\mu}{\sigma^2}$ بر روی ضرایب، ضرایب کوچک حذف می گردد. جهت تقویت ویژگی های بافتی تصویر، از یک عملگر ریخت شناسی ضخیم سازی به صورت مربعی با اندازه 3×3 استفاده می گردد و ماتریس تصاویر $DilatedH_1$ ، $DilatedV_1$ ، $DilatedH_2$ و $DilatedV_2$ شکل می گیرد. سپس تصویر ترکیبی حاصل از مجموع زیرباندهای فوق به صورت زیر تعریف می شود:

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

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 همسان سازی (نگاشت) هم اندازه می شوند. پس از همسان سازی ماتریسهای فوق، با قاعده زیر مقادیر باینری تصویر ناحیه بندی شده شکل می گیرد: به ازای هر درایه از ماتریس HybridCoef، اگر مقادیر متناظر آن در هر دو ماتریس $DilatedH_1$ و $DilatedV_1$ به طور همزمان صفر باشند، مقدار آن با $\min(DilatedH_1, DilatedV_1)$ جایگزین می گردد در غیر این صورت مقدار قبلی خود را حفظ می کند. در ادامه، جهت اتصال نواحی کوچک نزدیک به یک ناحیه بزرگتر، بار دیگر از یک عملگر ریخت شناسی ضخیم سازی به صورت مربعی با اندازه 4×4 استفاده شده و بر روی ماتریس ترکیبی HybridCoef اعمال می گردد. معمولاً، پس از اعمال عملگر ریخت شناسی ضخیم سازی بر روی تصویر ترکیبی، نواحی کوچک نویزی شکل می گیرد که با یک فیلتر ریخت شناسی با آستانه 300 پیکسل حذف می شوند. در انتها، با برجسب گذاری نواحی شکل گرفته بر روی تصویر ترکیبی، نواحی جدا شده در تصویر اصلی مشخص و تعیین می گردند. با مجموعه ترندهای فوق، مشکل چسبیدن نواحی متنی به گوشه های نواحی گرافیکی تصویر برطرف می گردد.

۲-۱-۲- ناحیه بندی مبتنی بر آستانه گذاری

پس از انجام ناحیه بندی اول یعنی ناحیه بندی مبتنی بر تبدیل موجک بر روی تصویر ورودی و حذف نواحی متنی شناسایی شده توسط این الگوریتم، الگوریتم ناحیه بندی دوم یعنی ناحیه بندی مبتنی بر آستانه گذاری بر روی باقی مانده تصویر اعمال می گردد. روند نمای الگوریتم ناحیه بندی مبتنی بر آستانه گذاری در شکل (۳) نشان داده شده است. در این الگوریتم، در ابتدا با پیدا کردن آستانه از روی هیستوگرام تصویر خاکستری، تصویر باینری می شود. سپس با عملگر ریخت شناسی ضخیم سازی به صورت مربعی با اندازه 20×20 ، نواحی کوچک نزدیک یکدیگر، به هم متصل می شوند.



شکل (۴): مراحل پیش‌پردازش و طبقه‌بندی نواحی
Fig. (4): Different stages of processing and segmentation

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. ۵۷۷۸ 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

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

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

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

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

طبقه‌بندی پیش‌پردازش‌شده‌ها 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.

۲-۱- 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 آنترپی مستخرج از ماتریس هم‌خدادی [۱۳] و همچنین دو ویژگی مستخرج از زیرباند های 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 هم‌خدادی، 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 3×3 دسته ضریب جزئیات در سه راستای افقی، عمودی و قطری وجود دارد بنابراین $2 \times 3 \times 3$ ویژگی شکل می‌گیرد. در این روش از دسته ضرایب تقریب برای استخراج ویژگی استفاده نمی‌شود.

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

۳-۲- طبقه‌بندی نواحی تصاویر

در انتهای هر مرحله ناحیه‌بندی با استخراج ویژگیهای تشریح‌شده، 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.

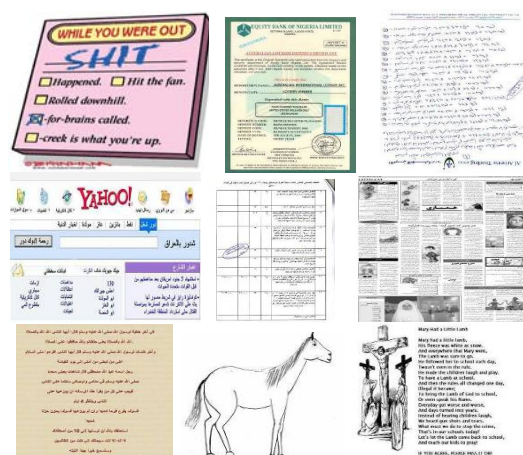
۳-۳- ناحیه‌بندی دومرحله‌ای

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 نزدیکی 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 سفیدرنگ انجام می‌گردد. تصویر خاکستری شده معکوس می‌گردد و پردازشها بر روی آن انجام می‌شود. نتایج میانی اعمال ناحیه‌بندی 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 ناحیه‌بندی 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 طبقه‌بندی اول از تصویر، ناحیه‌بندی 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 ناحیه‌بندی دوم نواحی عکسی با دقت بالایی از بقیه تصویر جدا می‌شوند. در انتها، نواحی شناسایی شده توسط الگوریتم ناحیه‌بندی دومرحله‌ای (متنی و عکسی) 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 به درستی تشخیص داده شده‌اند و محل‌های امضاء، لوگوهای موجود در تصاویر و جداول به عنوان نواحی غیرمتنی (نواحی شکلی) در نظر 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.

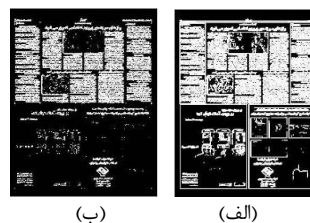


شکل (۵): نمونه‌ای از انواع مختلف تصاویر پایگاه داده
Fig. (5): A sample of different database images



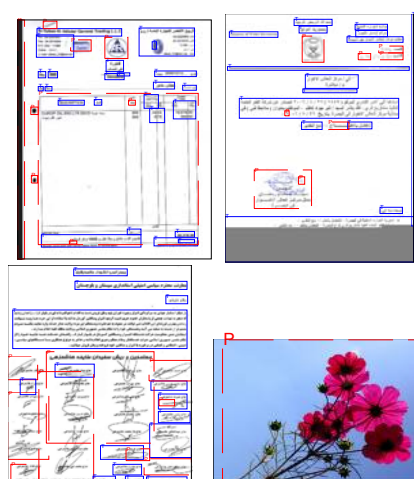
شکل (۶): تصویر نمونه روزنامه‌ای جهت بررسی نتایج میانی الگوریتم ناحیه‌بندی دومرحله‌ای. (الف): تصویر نمونه (ب): تصویر خاکستری معکوس شده.

Fig. (6): The newspaper sample image to investigate the two stage segmentation algorithm. (a): The sample image. (b): The reversed gray image



شکل (۷): نتایج اعمال ناحیه‌بندی 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 ناحیه‌بندی اول

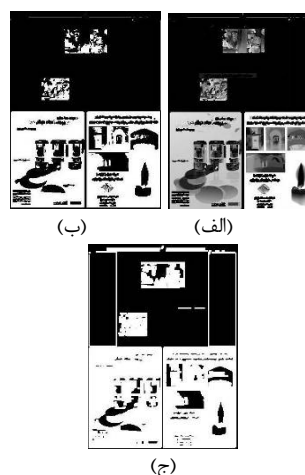
Fig. (7): The results of applying first stage of segmentation on the image. (a): The combined image resulting from overlapping the wavelets subbands. (b): The combined image after applying the threshold. (c): The final image of segmentation by first region contraing



شکل (۱۰): عملکرد الگوریتم پیشنهادی ناحیه بندی و طبقه بندی بر روی نمونه تصاویر. نواحی متنی با مستطیلهای آبی رنگ با خطوط ممتد و نواحی غیرمتنی با مستطیلهای قرمز رنگ با خطوط خطچین مشخص شده اند.
 Fig. (10): The effect of the proposed regional and segmentation algorithms on the sampled image. The texture regions are shown by blue rectangles and non-texture regions by dotted red lines

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٪ نسبت به بهترین نتایج الگوریتم‌های قبلی را نشان می‌دهد. این بهبود قابل توجه را می‌توان در استفاده مناسب و به‌جا از ساختار طبقه بندی سلسله‌مراتبی به همراه ویژگیهای خوب مورد استفاده قرار داد. علاوه بر این استفاده از یک الگوریتم ناحیه بندی دومرحله‌ای به همراه طبقه بندی نواحی و سپس تصمیم گیری 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 طبقه بندی 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 ناحیه بندی و همچنین سلسله‌مراتبی بودن طبقه بندیها جستجو کرد. اما در مقابل با استفاده از ویژگیهای متناسب با هر دسته از تصاویر و چندین طبقه بندی مختلف در یک ساختار سلسله‌مراتبی عملکرد طبقه بندی به صورت قابل توجهی بهبود می‌یابد.



شکل (۸): نتایج ناحیه بندی دوم بر روی باقی مانده تصویر حاصل از ناحیه بندی اول. (الف): حذف نواحی متنی از تصویر ورودی، (ب): تصویر باینری شده با استفاده از هیستوگرام رنگ، (ج): تصویر پس از عمل ضخیم سازی
 Fig. (8): The results of the second segmentation or the remaining of the first image of the first segmentation. (a)- elimination of the texture regions from the input image (b)- the binary image using the colour histogram (c)- the image after thickening action



شکل (۹): نواحی شناسایی شده توسط ناحیه بندی دومرحله‌ای
 Fig. (9): The identified regions by two steps segmentation

۳-۵- طبقه بندی نواحی حاصل از ناحیه بندی

در این قسمت عملکرد طبقه بندی سلسله‌مراتبی در طبقه بندی نواحی حاصل از الگوریتم ناحیه بندی دومرحله‌ای مورد بررسی قرار می‌گیرد. نتایج طبقه بندی نواحی شکلی، متنی و نواحی رد شده به صورت مجزا ارائه می‌گردد. نتایج طبقه بندی 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. طبقه بندیها 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 طبقه بندیها ۱۹۸۴۲ ناحیه را به درستی شناسایی کرده‌اند. در حالیکه ۵۰۳ ناحیه را نتوانسته‌اند به درستی شناسایی کنند. یعنی در مجموع صحت طبقه بندی برابر ۹۷.۵٪ بوده است و تنها ۲.۵٪ نواحی شناسایی نشده‌اند. البته به دلیل وجود ناحیه بندی دوم، مقدار FNR تأثیر چندانی در کارایی روش پیشنهادی ندارد. زیرا نواحی رد شده در مرحله اول توسط ناحیه بندی دوم مجدداً ناحیه بندی می‌شوند که باعث کاهش اثر نواحی رد شده در کارایی روش پیشنهادی می‌شود.

Table (1): The accuracy of classification of identified segments in 9 documents

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

تعداد سند	تعداد موارد صحیح (TA)	تعداد موارد اشتباه (FA)	جمع کل	بازیابی-دقت (%)
977	19842	503	20345	97.52

Table (2): Rejected regions by classifiers

جدول (۲): نواحی رد شده توسط طبقه‌بندها.

تعداد سند	تعداد موارد اشتباه (FR)	تعداد موارد صحیح (FA)	جمع کل	FA (%)	FR (%)
977	336	798	1134	70.4	29.6

Table (3): Comparison of the proposed algorithm and the previous algorithms

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

الگوریتم	طبقه‌بند مورد استفاده	ویژگی مورد استفاده	صحت طبقه‌بندی (%)
الگوریتم [۱۲]	SVM	هم‌خدادی و موجک	94.2
الگوریتم [۱۴]	SVM	موجک	91.4
الگوریتم [۱۵]	SVM	گردادیان	94.3
الگوریتم پیشنهادی	SVM & Neural network	موجک و هم‌خدادی	97.5

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

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 به وسیله 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 هم‌خدادی و همچنین دو ویژگی مستخرج از زیرباند های تبدیل موجک، جهت طبقه‌بندی و شناسایی برجسته نواحی شکل گرفته توسط الگوریتم ناحیه‌بندی استفاده می‌شود. طبقه‌بند سلسله‌مراتبی از دو طبقه‌بند پرسپترون چندلایه و یک طبقه‌بند ماشین بردار پشتیبان

References

- [1] A.K. Jain, B. Yu, "Document representation and its application to page decomposition", IEEE Trans. Patt. Anal. and Mac. Intel., Vol.20, pp.294-308, 1998.
- [2] G. Nagy, "Twenty years of document image analysis in PAMI", IEEE Trans. Patt. Anal. and Mach. Intel., 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. Intel., 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", Sys. and Cont. 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. Intel., Vol.15, pp.1162-1173, 1993.
- [9] G. Nagy, S. Seth, M. Viswanathan, "A prototype document image analysis system for technical journals", Computer, Vol.7, pp.10-22, 1992.
- [10] A. Jain, 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 Proc. 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 Sys., 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 Sys. for Vid. Tech., pp.256-268, 2002.

رزومه



حسین پورقاسم استادیار گروه مخابرات دانشکده برق دانشگاه آزاد اسلامی واحد نجف آباد، 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.